华章数学译丛

38

Real Analysis and Probability

(Second Edition)

实分析和概率论

(原书第2版)

(美) R. M. Dudley 麻省理工学院 著

赵选民 孙浩 译

机械工业出版社
China Machine Press

38

Real Analysis and Probability
(Second Edition)

实分析和概率论
(原书第2版)

（美） R. M. Dudley 著
麻省理工学院
赵选民 孙浩 译

机械工业出版社
China Machine Press

本书清晰地讲解了现代概率论以及概率测度与度量空间之间的相互关系．本书分两部分，第一部分介绍了实分析的内容，包括基础集合论、一般拓扑、测度、积分、巴拿赫空间及希尔伯特空间上的函数分析、凸集和函数以及拓扑空间上的测度，第二部分介绍了基于测度论上的概率论，包括大数定律、遍历定理、中心极限定理、条件期望、鞅收敛．另外，随机过程一章介绍了布朗运动以及布朗桥．

本书适合于概率论与数理统计方向的研究生，以及与之相关的研究生阅读，也适合于数学系高年级学生以及数学研究工作者参考使用．

本书版权登记号：图字：01-2006-2852

图书在版编目(CIP)数据

实分析与概率论(原书第2版)/(美)达德利(Dudley, R. M.)著；赵选民，孙浩译．—北京：机械工业出版社，2008.6

(华章数学译丛)

书名原文：Real Analysis and Probability, Second Edition

ISBN 978-7-111-23480-7

Ⅰ．实…　Ⅱ．①达…　②赵…　③孙…　Ⅲ．①实分析　②概率论　Ⅳ．O174.1　O211

中国版本图书馆 CIP 数据核字(2008)第054804号

机械工业出版社(北京市西城区百万庄大街22号　邮政编码　100037)

责任编辑：王春华

山西新华印业有限公司新华印刷分公司印刷·新华书店北京发行所发行

2008年6月第1版第1次印刷

186mm×240mm·24印张

标准书号：ISBN 978-111-23480-7

定价：55.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

本社购书热线：(010)68326294

译 者 序

2007 年 2 月 4 日，我尊敬的导师赵选民教授(译者之一)带着未能完成该书翻译的遗憾，永远地离开了爱着他的学生和他心爱的工作岗位．本书现在顺利出版，特以告慰赵老师的在天之灵．

本书主要包含两部分内容：实分析与概率论．这是一本非常系统的、完整的以实分析中的测度论理论为基础，揭示概率论的规律与理论的宏著．该论著主要适合于概率论与数理统计方向的研究生，以及与之相关方向的研究生阅读，也适合于数学系高年级学生以及数学研究工作者参考使用．

本书是在机械工业出版社华章分社大力支持下翻译完成的，翻译工作历时一年两个月．在翻译中，主要强调在尊重原文的基础上，数学专有名词翻译的规范性以及数学论证的逻辑性．译著经过多次讨论与修改，才形成现有状态．但由于译者自身知识水平有限，瑕疵在所难免，希望各位同仁不吝赐教，以便修正．

本书的翻译工作主要由赵选民教授和孙浩博士负责，刘向增、张燕、李颖、刘威、王彩玲、丛方媛、牛玉俊和何华等同志参与了讨论、校对、整理和编辑等相关工作，机械工业出版社华章分社的编辑们对译文提出了许多中肯的修订意见．在此向所有支持、参与本书出版的人士表示衷心的感谢，也将此译著献给 R. M. Dudley 教授，感谢他为我们系统学习测度论基础上的概率论提供了范本．

孙 浩

2008 年 4 月 25 日于西安

前　　言

本书是为研究生一年级学生而编写的．对欧几里得空间中的媒介分析的了解将对学习本书有很大的帮助，但是在本书中我并未将它列为预备知识．为了使这本书更加完整，我在第 1 章里增加了实数系和斯通–魏尔斯特拉斯定理，删除了一些其他书中已有的证明，并保留了像附录 B 中的复变理论方面的资料．

第 1 ~ 5 章提供的实分析可供一学期学习．之后的内容是可供一学期学习的概率论，这包括第 8 ~ 10 章，以及第 11 章、第 12 章的部分内容．标有 * 的段落和章节不常用到，至少是在以后很少用到，所以在第一次阅读时可以跳过，我们可以在以后用到时再读，这包括第 6 章中的一些和第 7 章的大部分内容．

一些不太重要的证明留给读者作为练习．如果读者发现本书的一些漏洞或者错误还请告知，我会很高兴地加以改进修正．虽然我认真地检查了所有的习题和提示，但是错误在所难免．提示也可能会起到误导的作用，所以请读者先尝试着自己做习题，尽可能不要看提示．

我尽自己最大的努力给出定理的最直接、最简明的证明．关于度量空间的完备化、强大数定律、遍历定理、鞅收敛定理、次加性遍历定理、Hartman-Wintner 重对数律的证明在相应的期刊中可以找到．

1950 年前后，当 Halmos 经典测度论问世的时候，测度论更高级的观点——局部紧空间也出现了，即本书 7.3 节中的内容．之后，概率论研究的主要方向转移到度量空间．第 11 章给出了较前沿的度量和概率之间联系的一些事实．附录 E 指出了在非度量（局部）紧空间中测度可能出错的地方．这部分内容仅仅经过一年的学习是不能很好地理解的，这需要一个长期的过程，但本书为眼前和以后的研究介绍了一些本质的内容．

在每章末都有习题，习题按照由简到难的顺序给出，并且多数习题都有解题提示，本版书中增加或改进了一些提示．

我也尽量给出书中定理的出处，有时这些历史注释和参考文献会很多，我都把它们罗列在各章的后面．有些注释被引申，有的被更正．但是，书中有关这些定理历史发展的评论仅是一家之言．

本书从 1967 年在麻省理工学院开始到 1976 年在丹麦奥尔胡斯完成．非常感谢 Ken. Alexander、Deborah Allinger、Laura Clemens、Ken Davidson、Don Davis、Persi Diaconis、Arnout Eikeboom、Sy Friedman、David Gillman、José Gonzalez、E. Griffor、Leonid Grinblat、Dominique Haughton、J. Hoffmann-Jørgensen、Arthur Mattuck、Jim Munkres、R. Proctor、Nick Reingold、Rae Shortt、Dorothy Maharam Stone、Evangelos Tabakis、Jin-Gen Yang，以及其他的同事和学生，感谢他们提出的宝贵建议．

在第 1 版出版时，Ken Brown 给出很多有益的建议，在此表示感谢，还要感谢 Justin Corvino、Charles Goldie、Charles Hadlock、Michael Jansson、Suman Majumdar、Rimas Norvaiša、Mark Pinsky、Andrew Rosalsky. Rae Shortt 和 Dewey Tucker. 特别需要感谢的是 Andries Lenstra、Valentin Petrov，他们给出了许多更好的建议．我们主要对 10.2 节的正则条件概率及第 12 章的马尔可夫时作了修改．

R. M. Dudley

目录

第 1 章　基础知识：集合论

在建造房子时，建设者会用与房子其他部分不一样的材料和方法来打造地基. 同样，几乎每一个数学分支都以公理集合论作为其基础. 这个基础被大多数关注基础理论的逻辑学家和数学家所接受，但是只有少数数学家有时间或意愿去详细研究公理集合论.

做另一个比喻，高级计算机语言和用它们编写的程序建立了计算机硬件和软件的基础. 但是，编写高级计算机程序的人需要了解多少计算机硬件和操作系统的知识则取决于他手边的问题.

在现代实分析中，集合论的问题比以前代数、复分析、几何和应用数学中的问题要多. 例如，在实分析中相对较近发展的“非标准分析”允许正数可以无限小但不为零. 非标准分析比早期发展的实分析更强地依赖于集合论的特性.

本章将介绍本书后面要用到的一些集合论符号和概念. 换句话说，本章只给出最基本的集合论知识. 附录 A 较详细地介绍了集合论，包括一些集合论公理，但是本书不打算介绍非标准分析或者更深入地讨论集合论.

本章所定义的许多概念在以后的章节里都要用到，希望读者能熟记.

1.1　集合论的定义和实数系

定义至少有两个目的. 首先，就像一本普通字典一样，定义试图给出见解，传达一种思想，或者用熟悉的概念去解释陌生的概念，但并不详尽说明或彻底研究所定义单词的全部意义. 我们称这 1
种定义为非形式(informal)定义. 在大多数数学和其他科学领域中，形式(formal)定义是完全准确的，因此，人们可以科学地判断一个有关命题的真伪. 在形式定义中，一个熟悉的术语(例如，普通的长度单位或数字)可以用不熟悉的术语定义. 集合论中的大多数定义都是形式的. 其次，集合论的另一个目的是不但为自己也为所有数学分支提供清晰的逻辑结构. 由此就产生从哪里开始定义的问题.

非形式的字典定义常常由一些同义词组成. 例如，字典中用“high”和“tall”相互定义，这种定义方法对于知道其中一个词的人来说是有帮助的，但是对于一个通过字典学习英语的人来说，这样的定义是无用的. 这种情况在一定程度上反映了人们在学习中所遇到的困难，因为字典中所有单词都是用其他单词来定义的. 因此，人们在开始时应至少弄清字典中一些单词的意思，而不是在用到时才去翻字典.

有些单词(例如，“and”、“or”和“but”这三个连词)非常常见，但很难用另外的词来定义. 这时，我们可能通过一些规则来定义含有连词的句子的含义，其中用这些连词连接起来的单词或短语的词义已经给定.

乍一想，你也许认为集合论中最重要的定义是“集合”，但是恰恰相反，因为数学的整个逻辑结构归约为集合或者由集合来定义，因此不能给出它的一个形式的、准确的定义. 相反，有一些规则(公理、推理规则等)能起到定义集合的作用. 一个初步的、非形式的集合(set)定义是“任何数学对象的全体”，但是随着学习的深入我们在对这个定义不断地加以明确和修正.

在某些方面，定义集合的问题类似于定义数(number)的问题. 经过几年的学习之后，学生们知

道了0，1，2，…这些数，并且知道了它们之间的运算规律．但是许多人还是很难确切地说明什么是数．即使人们完全同意算术规律，但是不同的人可能会给出数字1不同的定义.

在19世纪后期，数学家开始研究数的精确定义．一种方法是从0开始，通过取“后继者”或“下
2 一个更大的整数”的方法来产生更多的数.

如果定义了0，那么0之后的数字也就确定了．如果用n'表示n之后的数字，那么我们用0及它的后继者来定义一个序列0，$0'$，$0''$，$0'''$，…我们可以得到通常的整数定义．为了做到这一点，我们将在等号之前加上冒号“:=”表示“定义为相等”，例如$1:=0'$，$2:=0''$，$3:=0'''$，$4:=0''''$，等等．这样一直做下去，这些定义是精确的．这样，人们可以得到一本厚厚的数的字典，相当精确(尽管不太实用)但还不完整，因为0和后继运算没有被形式定义．大多数数系结构可以通过给定0和后继者的规则来确定．例如，一种规则是：假若$m'=n'$，则$m=n$.

一旦有足够多的规则来定义非负整数的结构，那么重要的是结构而不是结构中的每个具体的元素.

简言之，如果我们想要尽可能精确地建立严格的数学逻辑结构，那么非形式的定义不能是该结构的一部分，尽管在有些情况下，非正式的定义会帮助人们解释结构．然而，我们至少还要留下一些未被定义的基本概念．本节我们将给出公理体系和其他规则，并用基本概念定义其余的概念.

下面我们再次给出集合的非形式定义：一个集合是任何对象的全体．在数学上，这种对象将是数学对象，例如数、点、向量，或者其他的集合．(实际上，从集合论的观点来看，所有的数学对象都是一种集合或者另一种集合．)如果对象x是集合y中的元素，记作“$x\in y$”，读作“x属于y”或者“x在y中”．如果集合S是一个有限集，即S中的元素可以排成一个有限列x_1，x_2，…，x_n，那么S可以写作$S=\{x_1, x_2, \cdots, x_n\}$．例如，$\{2, 3\}$是只有2和3两个元素的集合．其中记号“$\in$”也是没有被正式定义的几个基本记号之一.

一个集合可以只有一个元素，例如，只有一个元素x的集合，记作$\{x\}$，读作“单元素集x”．在集合论中，$\{x\}$与x本身是不同的．例如，如果$x=\{1, 2\}$，则x有两个元素，而$\{x\}$只有一个元素.

集合A包含于(included)集合B中，或者A是B的子集(subset)，记作$A\subset B$，当且仅当A中的每一个元素都是集合B中的元素．一个等价的说法是，集合B包含(include)集合A，记作$B\supset A$．集合B包含(contain)x是指$x\in B$．当$B\supset A$，许多人也称B包含了A.

“if and only if”有时简记为“iff”．例如，$A\subset B$ iff 对所有x，如果$x\in A$，则$x\in B$.

集合论中一个重要的规则称为“外延性”．它是说，如果两个集合A和B有完全相同的元素，即
3 对任意的元素x，$x\in A$当且仅当$x\in B$，或者等价地$A\subset B$且$B\subset A$，那么两个集合相等，记作$A=B$．例如，$\{2, 3\}=\{3, 2\}$．集合中元素之间的次序可以是不同的，而它们的元素是相同的．在某种意义上，外延性是集合相等的一个定义．在集合论中另一种比较常用的观点是，任意两个对象是相等的当且仅当它们是完全同一的．因此，“$\{2, 3\}$”和“$\{3, 2\}$”是名字不同的同一个集合.

外延性也可以用来给出集合的非形式定义．一个集合可以仅以其所含的元素来定义，但元素之间的结构和关系与集合的定义无关.

与把有限个元素一一罗列出来不同，定义一个特殊集合的主要方法是给出元素所满足的条件．记号$\{x: \cdots\}$表示集合中的所有x满足…例如，集合$\{x: (x-4)^2=4\}=\{2, 6\}=\{6, 2\}$.

一种惯用的方法是在某一符号上划一道线表示“非”，例如，$a\neq b$表示“a不等于b”，符号“$\notin$”表示“不是…中的元素”．因此，$x\notin y$表示x不是y中的元素，例如$3\notin\{1, 2\}$.

由条件定义集合如果不仔细可能导致矛盾．例如，令 $r=\{x: x\notin x\}$，则 $r\notin r$ 意味着 $r\in r$，反之也成立(这就是罗素(Bertrand Russell)悖论)．这个悖论可以通过限制某些集合的条件而得以避免．因此，$\{x\in A: \cdots x\cdots\}$ 表示"A 中满足条件 $\cdots x\cdots$ 的所有 x 的集合"．当 A 是一个已知的集合时，只要使用这种形式的定义，用此方法就能定义新的集合，并且不会出现矛盾．

不管怎样，一个集合是它自己的元素看起来似乎有点奇怪，我们将在附录 A 中给予证明(定理 A.1.9)．从附录 A 给出的集合论的公理知，没有集合是它自己的元素．从这个意义上讲，在罗素悖论中命名的集合 r 就是所有集合的全体，在集合论中有时称之为"全域"．这里集合的非形式定义(即集合是任何对象的全体)确实是不严格的．附录 A 中的公理给出了一些条件，在该条件下某些全体可能是集合也可能不是集合．例如，全域不是一个集合．

数学上人们常常关心一个固定点集 y 的内部情况．我们用 $\{x: \cdots x\cdots\}$ 来表示 $\{x\in y: \cdots x\cdots\}$．

下面给出集合论中各种运算的定义．在这里我们不明确指明集合中的元素，公理就意味着它们是集合中的元素(见附录 A)．

没有任何元素的集合称为"空集"，记作 $\varnothing$，即对所有的 x，都有 $x\notin\varnothing$．由外延性，这个集合是唯一的．如果 B 是任意一个集合，则 2^B 称为 B 的"幂集"，是 B 的所有子集构成的集合．例如，如果集合 B 含有 3 个元素，则 2^B 有 $2^3=8$ 个元素．同时，$2^{\varnothing}=\{\varnothing\}\neq\varnothing$． 4

$A\cap B$ 称为 A 和 B 的交集，定义为 $A\cap B:=\{x\in A: x\in B\}$．换句话说，$A\cap B$ 是所有既属于 A 又属于 B 的 x 所组成的集合．$A\cup B$ 称为 A 和 B 的并集，表示对任意的 x，$x\in A\cup B$ 当且仅当 $x\in A$ 或 $x\in B$(或同时属于两者)．$A\setminus B$(读作"A 减 B")，表示所有在 A 中而不在 B 中的 x 所组成的集合，有时称为(B 在 A 中的)相对补集(relative complement)．对称差(symmetric difference) $A\Delta B$ 定义为 $(A\setminus B)\cup(B\setminus A)$．

用 $\mathbb{N}$ 表示所有非负整数 0，1，2，… 的集合．(形式上，非负整数通常通过定义 0 为空集 $\varnothing$，1 为 $\{\varnothing\}$ 来定义，一般地，后继的数由上面提到的运算定义 $n'=n\cup\{n\}$ 而得到，更详细的说明参见附录 A.)

有序对(ordered pair)非形式地定义为由一对给定次序的数学对象组成，例如 $\langle x, y\rangle$，这里 x 称为有序对 $\langle x, y\rangle$ 的第一个元素，y 为第二个元素．有序对满足下面的公理：对所有的 x，y，u 和 v，$\langle x, y\rangle=\langle u, v\rangle$ 当且仅当 $x=u$，$y=v$．在有序对 $\langle x, y\rangle$ 中可以出现 $x=y$ 的情况．有序对可以形式地通过满足公理的(无序的、平凡的)集合来定义．通常的方法是令 $\langle x, y\rangle:=\{\{x\}, \{x, y\}\}$(见附录 A)．由外延性知，$\{\{x\}, \{x, y\}\}=\{\{y, x\}, \{x\}\}$．

数学中的主要概念之一是函数．函数的非形式定义是，给定两个集合 D 和 E，D 上的函数定义为：对于 D 中的每一个元素 x，在 E 中有唯一的元素 $f(x)$ 和它对应．函数的形式定义是，一个函数(function)定义为有序对 $\langle x, y\rangle$ 的一个集合 f，使得对于任意的 x，y 和 z，如果 $\langle x, y\rangle\in f$ 且 $\langle x, z\rangle\in f$，则 $y=z$．例如，$\{\langle 2, 4\rangle, \langle -2, 4\rangle\}$ 是一个函数，但是 $\{\langle 4, 2\rangle, \langle 4, -2\rangle\}$ 不是函数．函数的有序对的集合非形式地称为函数的图形(graph)(对于 $D=E=\mathbb{R}$ 的情况，$\mathbb{R}$ 是实数集)．

函数 f 的定义域(domain)是使得对某些 y，$\langle x, y\rangle\in f$ 的所有 x 组成的集合，记作 $\operatorname{dom} f$．由函数的定义，y 是唯一确定的，记为 $f(x)$．f 的值域(range)是使得对于某些 x，$f(x)=y$ 的所有 y 的集合，记作 $\operatorname{ran} f$．

具有定义域 A 并且值域包含在集合 B 中的函数 f 称为是定义在 A 上的或是从 A 映射到 B 的．如果 f 的值域等于 B，则称 f 是映上(onto)到 B 的．

符号“$\mapsto$”有时用来表示或定义函数，一个函数可以写为“$x\mapsto f(x)$”. 例如，“$x\mapsto x^3$”或者“f: $x\mapsto x^3$”是指对(f的定义域中的)所有 x，使得 $f(x)=x^3$ 的函数f. 为明确定义域，通常使用与之相关的记号“f: $A\mapsto B$”，再加上f精确的定义，就表示它定义为从 A 映射到 B 的函数(而不是指
5 $f(A)=B$，为了区分记号 $\mapsto$的这两个应用，把 A 和 B 写成大写字母，把它们的元素写为小写字母，比如 x).

如果 X 是任意一个集合，A 是 X 的任意一个子集，A(在 X 上)的示性函数(indicator function)定义为

$$1_A(x):=\begin{cases}1, & 若\ x\in A\\ 0, & 若\ x\notin A.\end{cases}$$

(许多数学家称这个函数为 A 的特征函数(characteristic function.)在概率论中，“特征函数”是指傅里叶变换，这将在第 9 章介绍.)

一个序列(sequence)是定义在$\mathbb{N}$或者所有正整数集合$\{1, 2, 3, 4, \cdots\}$上的函数. 对所有的 n，满足$f(n)=x_n$ 的序列f通常写为$\{x_n\}_{n\geqslant 1}$.

在形式场合下，每个集合都是集合的集合(集合的每个元素也是一个集合). 如果一个集合被非形式地看作是由一些集合所组成的，那么我们称之为集族、集类，或者集合的全体. 设$\mathcal{V}$是一个集族，那么$\mathcal{V}$的并集(union)定义为

$$\bigcup\mathcal{V}:=\{x:对某些\ A\in\mathcal{V},x\in A\}.$$

同样，非空集族$\mathcal{V}$的交集(intersection)定义为

$$\bigcap\mathcal{V}:=\{x:对所有的\ A\in\mathcal{V},x\in A\}.$$

于是对任意两个集合 A 和 B，$\cup\{A, B\}=A\cup B$，$\cap\{A, B\}=A\cap B$. 诸如$\cup\mathcal{V}$和$\cap\mathcal{V}$的记号在集合论中是最常用的. 在数学的其他分支中，多于两个的集合的交集和并集经常用下标表示. 如果$\{A_n\}_{n\geqslant 1}$是一个集合序列，则它们的并集写为

$$\bigcup_n A_n:=\bigcup_{n=1}^{\infty}A_n:=\{x:对某些\ n,x\in A_n\}.$$

同样地，它们的交集写为

$$\bigcap_{n\geqslant 1}A_n:=\bigcap_{n=1}^{\infty}A_n:=\{x:对所有的\ n,x\in A_n\}.$$

有限个集合 A_1，$\cdots$，A_n 的并集写为

$$\bigcup_{1\leqslant i\leqslant n}A_i:=\bigcup_{i=1}^{n}A_i:=\{x:对某些\ i=1,\cdots,n,x\in A_i\}.$$

6 类似地，将并改为交，“一些”改为“所有的”，即可定义有限个集合的交集.

更一般地，令 I 是任意一个集合，设 A 是定义在 I 上的函数，其值是集合 $A_i:=A(i)$，那么所有这些集合 A_i 的并集写为

$$\bigcup_i A_i:=\bigcup_{i\in I}A_i:=\{x:对某些\ i,x\in A_i\}.$$

在这种情况下，集合 I 称为指标集(index set)，这意味着 I 是函数 $i\mapsto A_i$ 的定义域. 如果从上下文可以清楚地看出 I 是什么，就像上面的第一个表示式那样，指标集在记号中可以被省略. 同样地，交集可以写为

$$\bigcap_i A_i := \bigcap_{i\in I} A_i := \{x: \text{对所有的 } i \in I, x \in A_i\},$$

这里的 I 通常是一个非空集合．但当这些集合是所讨论的某一固定集合（比如说 X）的一切子集时，就出现例外．假定 $t \notin I$ 并且令 $A_t := X$，如果 I 是非空集合，那么用 $I \cup \{t\}$ 代替 I 不会改变 $\bigcap_{i\in I} A_i$. 在 I 是空集的情况下，我们可以令 $\bigcap_{i\in\varnothing} A_i = X$.

下面介绍在数理逻辑中两个用得比较多的缩写符号：$\forall$ 表示"对任意的"，而 $\exists$ 表示"存在"．例如，$(\forall x\in A)(\exists y\in B)\cdots$ 表示对任意的 $x\in A$，在 B 中存在一个 y，使得…．

两个集合 A，B 称为*不相交的*（disjoint），当且仅当 $A\cap B=\varnothing$. 集合 A_i 对 $i\in I$ 称为是不相交的，当且仅当对 I 中所有的 $i\neq j$，$A_i\cap A_j=\varnothing$.

下面将给出不同数域的定义，首先是实数域．假定所有的读者都熟悉整数和有理数．更详细的介绍参见附录 A. 4.

回忆以前学过的：$\mathbb{N}$ 表示所有非负整数 0，1，2，…组成的集合，$\mathbb{Z}$ 表示所有整数 0，±1，±2，…组成的集合，$\mathbb{Q}$ 表示所有有理数 m/n 组成的集合，其中 $m\in\mathbb{Z}$，$n\in\mathbb{Z}$，并且 $n\neq0$.

实数可以用不同的方法来定义．一种常用的方法是用十进制数表示：x 是实数当且仅当 $x=\pm y$，其中 $y=n+\sum_{j=1}^{\infty} d_j/10^j\,(n\in\mathbb{N})$，并且每个数字 d_j 是 0～9 之间的整数．但是这种表示在分析证明中用起来不方便，而且对形如 $m/10^k\,(m\in\mathbb{Z},\ m\neq0,\ k\in\mathbb{N})$ 的有理数，表示是不唯一的. 还可以根据一般的收敛到实数 x 的有理数列来定义实数 x，当讲完度量空间后，我们将在 2. 5 节研究这一问题.

这里所用的实数的形式定义是用戴德金（Dedekind）分割法来定义，定义如下：一个*分割*（cut）是一个集合 $C\subset\mathbb{Q}$，使得 $C\notin\varnothing$；$C\neq\mathbb{Q}$；无论何时 $q\in C$，如果 $r\in\mathbb{Q}$，且 $r<q$，则 $r\in C$，并且存在 $s\in\mathbb{Q}$，满足 $s>q$ 且 $s\in C$. 7

令 $\mathbb{R}$ 表示所有实数的集合．那么形式上 $\mathbb{R}$ 就是所有分割的集合．也可非形式地将实数 x 与分割 C 之间的一一对应记作 $C_x=\{q\in\mathbb{Q}: q<x\}$，其中 $C=C_x$ 或者 $x=x_C$.

实数中的次序 $x\leqslant y$，根据分割定义为 $C_x\subset C_y$. 一个实数集 E 称为*上有界*（bounded above），它的*上界*（upper bound）为 y，当且仅当对任意的 $x\in E$，有 $x\leqslant y$. 称 y 为 E 的*上确界*（supremum）或者*最小上界*（least upper bound），记作 $y=\sup E$，当且仅当 y 是一个上界并且对 E 的任意一个上界 z，有 $y\leqslant z$. 关于 $\mathbb{R}$ 的一个基本事实是对每一个非空集 $E\subset\mathbb{R}$，使得 E 是上有界的，则上确界 $y=\sup E$ 存在．这用分割法很容易证明：对所有的 $x\in E$，令 C_y 是 C_x 的并集，证明见附录 A 的定理 A. 4. 1.

类似地，一个实数集 F 称为*下有界*（bounded below），它的*下界*（lower bound）为 v，如果对任意 $x\in F$，有 $v\leqslant x$，并且称 v 为 F 的*下确界*（infimum），记作 $v=\inf F$，当且仅当对 F 的任意一个下界 t，有 $t\leqslant v$. 每一个下有界的非空集 F 有下确界，即（非空且上有界的）F 的下界中的上确界.

两个实数的最大值和最小值定义为：若 $x\leqslant y$，则 $\min(x,y)=x$，$\max(x,y)=y$；否则有 $\min(x,y)=y$，$\max(x,y)=x$.

对任意的实数 $a\leqslant b$. 令 $[a,b]:=\{x\in\mathbb{R},\ a\leqslant x\leqslant b\}$.

对于任意的两个集合 X 和 Y，它们的*笛卡儿积*（Cartesian product），记作 $X\times Y$，定义为由 X 中的 x 和 Y 中的 y 构成的所有有序对 $\langle x, y\rangle$ 的集合．笛卡儿积的一个典型的例子是 $\mathbb{R}\times\mathbb{R}$，也可以写作 $\mathbb{R}^2$（读作 r-2，而不是 r-平方），称为*平面*（plane）.

习题

1. 设 $A:=\{3,4,5\}$, $B:=\{5,6,7\}$. 计算:(a)$A\cup B$.(b)$A\cap B$.(c)$A\setminus B$.(d)$A\Delta B$.
2. 证明 $\varnothing\neq\{\varnothing\}$, $\{\varnothing\}\neq\{\{\varnothing\}\}$.
3. 下列三组集合哪些是相等的?
(a)$\{\{2,3\},\{4\}\}$.(b)$\{\{4\},\{2,3\}\}$.(c)$\{\{4\},\{3,2\}\}$.
4. 下列哪些是函数,为什么?
(a)$\{\langle 1,2\rangle,\langle 2,3\rangle,\langle 3,1\rangle\}$.
(b)$\{\langle 1,2\rangle,\langle 2,3\rangle,\langle 2,1\rangle\}$.
(c)$\{\langle 2,1\rangle,\langle 3,1\rangle,\langle 1,2\rangle\}$.
(d)$\{\langle x,y\rangle\in\mathbb{R}^2: x=y^2\}$.
(e)$\{\langle x,y\rangle\in\mathbb{R}^2: y=x^2\}$.

8 5. 对于任意关系 V(即有序对所组成的任意集合),定义 V 的定义域为 $\{x$:对某些 y, $\langle x,y\rangle\in V\}$, V 的值域为 $\{y$:对某些 x, $\langle x,y\rangle\in V\}$. 求上题中各个关系的定义域和值域.(它们是否是函数?)
6. 设 $A_{1j}:=\mathbb{R}\times[j-1,j]$, $A_{2j}:=[j-1,j]\times\mathbb{R}$, 其中 $j=1,2$. 设 $B:=\bigcup_{m=1}^{2}\bigcap_{n=1}^{2}A_{mn}$, $C:=\bigcap_{n=1}^{2}\bigcup_{m=1}^{2}A_{mn}$, 则下面关系哪一个是正确的: $B\subset C$ 还是 $C\subset B$? 为什么?
7. 令 $f(x):=\sin x$, $x\in\mathbb{R}$. 对于$\mathbb{R}$的下列子集, f 映射到哪一个, f 映上到哪一个?
(a)$[-2,2]$.(b)$[0,1]$.(c)$[-1,1]$.(d)$[-\pi,\pi]$.
8. 如果将习题 7 中的弧度转换为度,会有什么结果?
9. 在下面的集合中,哪些包含于其他的集合中?
$$A:=\{3,4,5\};\ B:=\{\{3,4\},5\};$$
$$C:=\{5,4\};\ D:=\{\{4,5\}\}.$$
假定有明显的关系成立,例如 $4=\{3,5\}$. 更具体一点,假定对于任意两个集合 x 和 y, 下列三种关系至多有一个成立: $x\in y$, $x=y$, 或 $y\in x$, 并且每个非负整数 k 是一个有 k 个元素的集合. 请解释上述的关系中为何包含,为何不包含. 例如,如果 $\{\{6,7\},\{5\}\}\subset\{3,4\}$, 则由延伸性, $\{6,7\}=3$ 或 4, 但是 $\{6,7\}$ 只有两个元素,而不是三个或四个.
10. 令 $I:=[0,1]$, 计算 $\bigcup_{x\in I}[x,2]$, $\bigcap_{x\in I}[x,2]$.
11. "半闭直线"是实数集$\mathbb{R}$的子集,具有形式 $\{x\in\mathbb{R}: x\leqslant b\}$ 或 $\{x\in\mathbb{R}: x\geqslant b\}$, 其中的 b 是实数. $\mathbb{R}$ 上的 n 次多项式(polynomial of degree n)是一个函数
$$x\mapsto a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\quad 且\quad a_n\neq 0.$$
证明:当 n 为奇数时,任意 $n\geqslant 1$ 次多项式的值域是$\mathbb{R}$;而当 n 是偶数时,其值域是半闭直线.[提示:证明当 $|x|$ 充分大时,多项式与它的首项 a_nx^n 的符号相同,并且它的绝对值趋于∞. 对连续函数(如多项式)使用介值定理(参见习题 2.2.14(d)).]
12. $\mathbb{R}^2$ 上的多项式是具有如下形式的函数: $\langle x,y\rangle\mapsto\sum_{0\leqslant i\leqslant k,0\leqslant j\leqslant k}a_{ij}x^iy^j$. 证明: $\mathbb{R}^2$ 上非常数多项式的值域或者是$\mathbb{R}$, 或者是半闭直线,或者是半开直线 $(b,\infty):=\{x\in\mathbb{R}: x>b\}$ 或 $(-\infty,b):=\{x\in\mathbb{R}: x<b\}$, 这里每一个开的或闭的半直线是某个多项式的值域.[提示:对一半开直线,讨论多项式 $x^2+(xy-1)^2$ 的值域.]

1.2 关系和序

一个关系(relation)就是有序对的任意集合. 对于任意的关系 E, 逆关系定义为 $E^{-1}:=\{\langle y,x\rangle:$
9 $\langle x,y\rangle\in E\}$. 因此,函数是一种特殊的关系,它的逆 f^{-1} 不必是一个函数. 事实上,一个函数称为

1－1的或一对一的(one-to-one)，当且仅当f^{-1}也是一个函数．给定关系E，经常把$\langle x, y\rangle \in E$表示为$xEy$(注意：这种记号不用于函数而用于其他关系，我们将在后面解释)．给定一个集合X，关系$E \subset X \times X$称为在X上是自反的(reflexive)，当且仅当对所有的$x \in X$，有xEx．关系E称为对称的(symmetric)，当且仅当$E = E^{-1}$．称关系E是传递的(transitive)，当且仅当若xEy，yEz，则有xEz．一个具有传递关系的例子是序，例如$x \leqslant y$．

一个关系$E \subset X \times X$称为等价关系(equivalence relation)，当且仅当它在X上是自反的、对称的和传递的．等价关系的一个例子是等式．一般地，一个等价关系相当于一个等式．满足等价关系的两个元素x和y，在某种意义下是相等的．例如，两个整数m和n称为在模p的意义下相等，当且仅当$m-n$能被p整除．在模p的意义下相等是一种等价关系．或者如果f是一个函数，我们可以通过$xE_f y$定义等价关系E_f，当且仅当$f(x)=f(y)$．

给定一个等价关系E，一个等价类(equivalence class)是具有如下形式的集合：对任意$x \in X$，$\{y \in X: yEx\}$．从等价关系的定义推出，两个等价类或者不相交或者相等．设$f(x) := \{y \in X: yEx\}$，则f是一个函数且xEy，当且仅当$f(x)=f(y)$，因此$E = E_f$，并且每一个等价关系可以写成E_f形式．

一个关系E称为是反对称的(antisymmetric)，当且仅当若xEy且yEx，则$x=y$．给定一个集合X，一个偏序(partial ordering)是一个传递的、反对称的关系$E \subset X \times X$，那么$\langle X, E\rangle$称为偏序集(partially ordered set)．例如，对于任意一个集合Y，令$X = 2^Y$(Y的所有子集的集合)，则对于通常的包含关系$\subset$，$\langle 2^Y, \subset\rangle$构成了一个偏序集．(注：许多人认为偏序也是自反的，这里定义的偏序不但包括$\leqslant$，而且还有$<$．)一个偏序称为是严格的(strict)，如果xEx对于任意的x都不成立．因此"严格的"和"自反的"是相对的．对于任意的偏序集E，由$x \leqslant y$定义关系$\leqslant$，当且仅当(xEy或者$x=y$)，则$\leqslant$是一个自反偏序．由$x<y$定义关系$<$，当且仅当(xEy且$x \neq y$)，则$<$是一个严格偏序．例如，在实数之间通常的关系$\leqslant$和$<$就是上面所定义的偏序和严格偏序．$F = E \cup D$和$E = F \setminus D$定义了集合X的严格偏序E和自反偏序F之间的一个一一对应，这里D是"对角的"，$D := \{\langle x, x\rangle: x \in X\}$．从现在开始所考虑的偏序集或者是自反的，通常记作$\leqslant$或$\geqslant$；或者是严格的，记作$<$(或$>$)．这里，"$<$"通常读作"小于"，等等．

两个偏序集$\langle X, E\rangle$和$\langle Y, G\rangle$称为是有序同构的(order-isomorphic)，当且仅当存在一个从X映射到Y的一一对应(1－1且映上)的函数f，使得对任意的x，$u \in X$，有uEx，当且仅当$f(u)Gf(x)$． 10
那么，f称为有序同构(order-isomorphism)．例如，区间$[0, 1]$和$[0, 2]$，具有通常的实数之间的有序关系，通过函数$f(x)=2x$，它们是有序同构的．但是区间$[0, 1]$不与$\mathbb{R}$有序同构，它没有最小元素．

从现在开始，有序对$\langle x, y\rangle$一般记作(x, y)(这当然与无序对$\{x, y\}$是不一样的)．

集合X的线性序(linear ordering)E是X的偏序E，若$\forall x$，$y \in X$，有xEy，yEx，或者$x=y$，则称$\langle X, E\rangle$为线性序集(linearly ordered set)．线性序集的一个经典例子是实直线$\mathbb{R}$，满足通常的有序关系．实际上$(\mathbb{R}, <)$、$(\mathbb{R}, \leqslant)$、$(\mathbb{R}, >)$和$(\mathbb{R}, \geqslant)$都是线性序集．

如果$\langle X, E\rangle$是任意一个偏序集，A是X的任意一个子集，则$\{\langle x, y\rangle \in E: x \in A, y \in A\}$也是$A$上的一个偏序，记作$E_A$．对于大多数的序，就像实数上的序那样，子集的序与整个集合上的序用同一记号来表示．如果(X, E)是线性序的且$A \subset X$，则直接从定义知，(A, E_A)也是线性序的．

设W是一个具有自反线性序$\leqslant$的集合．W称为是良序的(well-ordered)，当且仅当对于W的任意非空集合A，在A中存在最小元素x，使得对A中任意的元素y，有$x \leqslant y$．相应的严格线性序$<$也

是一个良序．如果 X 是一个有限集，则 X 中的任意一个线性序都是良序．虽然区间 $[0, 1]$ 有最小元素 0，但它不是良序的，因为它有没有最小元素的子集，比如 $\{x: 0<x\leqslant 1\}$.

在良序集的证明中可以用数学归纳法．假设 $(X, <)$ 是良序集，我们想要证明 X 中的所有元素都具有某些性质．如若不然，则必有一个最小的元素不具有该性质．因此只需证明对每个 $x\in X$，若对所有 $y<x$，该性质都成立，则对 x 它也成立．这种"归纳原理"将在 1.3 节详细讨论.

习题

1. 对任意偏序 E，证明 E^{-1} 也是偏序．
2. 对两个偏序集 $\langle A, \leqslant\rangle$ 和 $\langle B, \leqslant\rangle$，笛卡儿积 $A\times B$ 的字典序(lexicographical ordering)如下定义：$\langle a, b\rangle\leqslant\langle c, d\rangle$，当且仅当 $a<c$ 或者 $(a=c, b\leqslant d)$.（例如，如果 A 和 B 两者都是具有通常次序的字母，则我们有
11 两个字母的单词或者字母串的字典序．）如果 A 与 B 上的序是线性的，证明：$A\times B$ 上的字典序也是线性的．如果通过给定的关系，A 和 B 是良序的，证明：$A\times B$ 也是良序的.
3. 设 $\langle a, b\rangle\leqslant\langle c, d\rangle$，当且仅当 $a\leqslant c$ 且 $b\leqslant d$，证明：这是一个偏序．如果 A 和 B 都包含不止一个元素，证明：这个序不是 $A\times B$ 上的线性序．
4. 在 $\mathbb{R}^2$ 上，设 $\langle x, y\rangle E\langle u, v\rangle$ 当且仅当 $x+y=u+v$，$\langle x, y\rangle F\langle u, v\rangle$ 当且仅当 $x+y\leqslant u+v$，$\langle x, y\rangle G\langle u, v\rangle$ 当且仅当 $x+u\leqslant y+v$. 则 E，F 和 G 中哪一个是等价关系，哪一个是偏序或者线性序？为什么？
5. 对实数列 $\{x_n\}$，令 $\{x_n\}E\{y_n\}$ 当且仅当 $\lim\limits_{n\to\infty}x_n-y_n=0$，而 $\{x_n\}F\{y_n\}$ 当且仅当 $\lim\limits_{n\to\infty}x_n-y_n=1$，则 E 和 F 中，哪一个是等价关系和(或)偏序？为什么？
6. 对于同一集合 X 上的任意两个关系 E 和 F，用 xGz 定义一个关系 $G:=E\circ F$，当且仅当对某个 y，xEy 和 yFz. 对下面的每一个性质：(a)自反的，(b)对称的，(c)传递的．如果 E 和 F 都具有该性质，证明：G 也具有该性质，若不具有则举出反例．
7. 根据习题 6，对如下性质回答同一问题.（a)反对称的，(b)等价关系，(c)函数.

*1.3 超限归纳和递归

数学归纳法是众所周知的并且是在证明与非负整数有关的命题时常用的方法．如果 $F(n)$ 是我们需要证明的对所有 $n\in\mathbb{N}$ 成立的命题，而直接的证明又不那样容易，那么先证 $F(0)$ 成立，再假定 $F(n)$ 成立，然后证明 $F(n+1)$ 也成立，这种方法往往是非常有帮助的．或者，也可以假设 $F(0)$，$F(1)$，…，$F(n)$ 都成立，然后证明 $F(n+1)$ 也成立．更一般地，设 $(X, <)$ 是任意一个偏序集，子集 $Y\subset X$ 称为*归纳的*(inductive)，如果对每一个 $x\in X$，使得 $y\in Y$，对所有 $y\in X$，使得 $y<x$，就有 $x\in Y$. 假如 X 有一个最小元素 x，则不存在 $y<x$，因此，x 一定属于 X 的任意一个归纳子集 Y. 在一般的归纳法中，设 Y 是所有使得 $F(n)$ 成立的 n 的集合．证明 Y 是归纳的就证明了 $Y=\mathbb{N}$，于是 $F(n)$ 对所有的自然数 n 都成立．在实数集 $\mathbb{R}$ 中，集合 $(-\infty, 0)$ 是归纳的，但它不是所有的 $\mathbb{R}$．集合 $\mathbb{N}$ 是一个良序集，但 $\mathbb{R}$ 不是，因为集合 $\{x\in\mathbb{R}: x>1\}$ 没有最小元素．良序集的主要优点之一是它们允许如下对归纳法的扩展.

12 **1.3.1 归纳原理** *设 X 是任意关于关系 $<$ 的良序集，Y 是 X 的任意归纳子集，则 $Y=X$.*

证明 如果 $X\setminus Y=\varnothing$，则结论显然成立.

若不然，设 y 是 $X\setminus Y$ 的最小元素，则对所有使得 $x<y$ 的 x，都有 $x\in Y$(如果 y 不是 X 的最小元素，显然成立)，因此，$y\in Y$，产生矛盾. □

对于任意线性序集 $(X, \leqslant)$，*初始段*(initial segment) $Y\subset X$ 是满足如下条件的集合：当 $y\in Y$，$x<y$ 时，有 $x\in Y$. 如果 $(X, \leqslant)$ 是有通常序的实直线，Y 是初始段，则要么对某些 y 有 $Y=X$ 成立，

要么 $Y=\{x: x<y\}$，或者 $Y=\{x: x\leqslant y\}$ 成立.

在通常的数学归纳法中，集合 $(X, <)$ 与非负整数的集合 $\mathbb{N}$ 是有序同构的，或者与 $\mathbb{N}$ 的初始段（有限整数）关于通常序是有序同构的. 超限归纳法是指对具有比较复杂的良序 $(x, <)$ 的归纳法. 例如“双归纳法”. 为了证明命题 $F(m, n)$ 对所有的非负整数 m 和 n 都成立，首先证明 $F(0, 0)$ 成立. 在证明 $F(m, n)$ 时，我们可以先假设 $F(j, k)$ 对于所有的 $j<m$ 和 $k\in\mathbb{N}$，及对 $j=m$ 和 $k<n$ 都成立.（在这种情况下，良序就是 1.2 节习题 2 中提到的字典序.）$\mathbb{N}\times\mathbb{N}$ 的其他良序也是有用的. 在集合论中，我们更关心的是一般的良序集而不仅是序列，比如 $\mathbb{R}$ 的良序，尽管这些没有本质意义上的差别.（一般集合的良序，特别是 $\mathbb{R}$ 的良序，依赖于选择公理，我们将在 1.5 节讨论.）

下面将给出数学中的另外一种重要的方法——递归. 在它的传统形式中，函数 f 可以由初始 $f(0)$，然后用 $f(n-1)$ 以及对 $k<n$ 的 $f(k)$ 其他可能的值确定 $f(n)$ 来定义. 这种递归定义可以推广到良序集上. 对于 $A\subset \operatorname{dom} f$ 和任意函数 f，函数 f 在 A 上的限制定义为 $f\restriction A:=\{\langle x, f(x)\rangle: x\in A\}$.

1.3.2 递归原理 *设 $(X, <)$ 是良序集，Y 是任意集合. 对任意 $x\in X$，设 $I(x):=\{u\in X: u<x\}$. 设函数 g 的定义域是所有满足如下条件的 j 的集合：对某些 $x\in X$，j 是从 $I(x)$ 映射到 Y 的函数，并且 $\operatorname{ran} g\subset Y$，则存在唯一的从 X 映射到 Y 的函数 f，使得对每个 $x\in X$，$f(x)=g(f\restriction I(x))$.*

注：*如果 b 是 X 中的最小元素，要想定义 $f(b)=c$，则只要让 $g(\varnothing)=c$，$I(b)=\varnothing$ 即可.*

证明 如果 $X=\varnothing$，那么 $f=\varnothing$，结论成立. 设 X 是非空集合，且 b 是它的最小元素. 对每个 $x\in X$，令 $J(x):=\{u\in X: u\leqslant x\}$，令 T 是所有满足如下条件的 $x\in X$ 的集合：在 $J(x)$ 上有一个函数 13
f，满足对所有 $u\in J(x)$，$f(u)=g(f\restriction I(u))$. 下面证明：如果这样的函数 f 存在，则它是唯一的. 设 h 是另外一个这样的函数，则 $h(b)=g(\varnothing)=f(b)$. 由归纳原理(1.3.1)可得，对每个 $u\in J(x)$，$h(u)=g(f\restriction I(u))=f(u)$. 因此 f 是唯一的. 如果对 T 中的某个 u，$x<u$，f 如上是定义在 $J(u)$ 上的函数，则 $f\restriction J(x)$ 有所期望的性质，并且由唯一性知，对 $J(x)$ 来说，它就是 f. 因此，T 是 X 的一个初始段. 对所有的 $x\in T$，所有这样的函数 f 的并是一个确定的函数，我们把它也称为 f. 如果 $T\neq X$，设 u 是 $X\setminus T$ 的最小元素. 那么 $T=I(u)$，且 $f\cup\{\langle u, g(f)\rangle\}$ 是 $J(u)$ 上的具有所期望性质的函数，因此 $u\in T$，矛盾. 所以函数 f 存在，因为它在每一个 $J(x)$ 上都是唯一的，所以它是唯一的. □

对笛卡儿积 $A\times B$ 上的任意函数 f，我们通常记作 $f(a, b)$ 而不是 $f(\langle a, b\rangle)$. 在非负整数上传统的递归阐述如下.

1.3.3 推论（简单递归） *设 Y 是任意集合，$c\in Y$，h 是从 $\mathbb{N}\times Y$ 到 Y 的任意函数，则存在从 $\mathbb{N}$ 到 Y 的函数 f，使得 $f(0)=c$ 且对每个 $n\in\mathbb{N}$，$f(n+1)=h(n, f(n))$.*

证明 为了应用递归原理 1.3.2，设 $g(\varnothing)=c$，令 j 是从某个非空集合 $I(n)$ 映射到 Y 的任意函数.（注意：$I(n)$ 是空集当且仅当 $n=0$）. 则 $n-1$ 是 $I(n)$ 中最大的数，设 $g(j)=h(n-1, j(n-1))$，则函数 g 由这样的函数 j 所完全确定，应用递归原理 1.3.2 得到函数 f，此时 $f(0)=c$，并且对任意 $n\in\mathbb{N}$，$f(n+1)=g(f\restriction I(n+1))=h(n, f(n))$. □

例：设 t 是 $\mathbb{N}$ 上的实值函数，令

$$f(n)=\sum_{j=0}^{n} t(j).$$

为了利用简单递归 1.3.3 得到 f，对于任意的 $n\in\mathbb{N}$，$y\in\mathbb{R}$，令 $c=t(0)$ 和 $h(n, y)=t(n+1)+y$. 给出 t 的算法程序后，根据递归算法就可以写出计算 f 的计算机程序，从而在某种意义下把求和归约为简单的加法.

例：一般递归原理 1.3.2 可以用来定义一个函数 f，使得对 $n=1, 2, \cdots$，$f(n)$是第 n 个素数：$f(1)=2$，$f(2)=3$，$f(3)=5$，$f(4)=7$，$f(5)=11$，等等. 在空函数上，函数 g 定义为 2，因此
14 $f(1)=2$. 在 $J(n)=\{1, 2, \cdots, n\}=I(n+1)$上给定 j，令 $g(j)$是使得 $k>j(n)$且 k 不能被$j(1)$，$j(2)$，…，$j(n)$整除的最小 k 值. 因为有无限多个素数，所以这样的 k 是存在的. 因此$f(1)<f(2)<\cdots$. 在构造函数f时，g 只用于具有性质 $j(1)<j(2)\cdots<j(n)$的函数 j（因为 $j(i)=f(i)$）. 注意到，在 $f(n+1)=g(j)$的定义中使用的是所有的 $j(1)$，…，$j(n)$的值而不仅仅是 $j(n)$的值.

习题

1. 为了用简单递归法定义阶乘函数$f(n)=n!$，如何选择推论 1.3.3 中的 c 和 h? 为使 h 达到这个目的，对哪些 n 和 x（如果任意），$h(n, x)$的值是唯一确定的?
2. 设$(X, <)$是一个良序集，$>$与通常的定义一样，即 $x>y$ 意味着 $y<x$，设$(X, >)$也是一个良序集，且 f 是从$\mathbb{N}$映上到 X 的函数，证明：f 不是一对一的.［提示：如果是，则在$\mathbb{N}$上通过递归定义函数 h，使得 $h(0)$对$<$是 X 的最小元素，$h(1)$是次小的元素，等等，则 h 的值域中没有最大元素.］
3. 给定一个偏序集$(X, \leqslant)$，子集 $A\subset X$，元素 $x\in A$ 称为 A 的极小(minimal)元素，当且仅当不存在 $y\in A$，使得 $y<x$. $(X, \leqslant)$称为极小序(min-ordered)当且仅当它的每一个非空子集有一个极小元素.（注意：任意良序集都是极小序集.）
   (a)证明：将归纳原理 1.3.1 中的良序换为极小序，原理仍然成立.
   (b)同样递归原理 1.3.2 也成立.
4. 对于习题 3，在$\mathbb{N}\times\mathbb{N}$上由$(i, k)\leqslant(m, n)$定义序当且仅当 $i\leqslant m$，$k\leqslant n$，
   (a) 证明：$(\mathbb{N}\times\mathbb{N}, \leqslant)$是极小序.
   (b)设 $A:=\{(m, n)\in\mathbb{N}\times\mathbb{N}: m+n\geqslant 4\}$，则 A 的极小元素是什么? 证明：$(\mathbb{N}\times\mathbb{N}, \leqslant)$不是良序.
   注：在集合论基础上建立数学时，我们已经定义了$\mathbb{N}$，通过$\mathbb{N}\times\mathbb{N}$上的字典序或者如上所定义的极小序的递归，可以给出加法和乘法的定义（参见附录 A.3 及其参考文献）.
5. 再一次对习题 3，设$(X, \leqslant)$是偏序集，其中 X 的唯一归纳子集是 X（像归纳原理中那样），证
15 明：$(X, \leqslant)$一定是极小序.
6. 设$(X, \leqslant)$是一个偏序集，在其上定义一个函数 f，使得对每一个 $x\in X$，有 $f(x)<x$，利用递归原理证明：存在从$\mathbb{N}$映射到 X 的函数 g，使得对所有的 n，$g(n+1)<g(n)$.

1.4 势

两个集合 X 和 Y 有相同的势(same cardinality)，当且仅当它们之间存在一个从 X 映射到 Y 的一一对应的函数. 我们说这两个集合的“元素个数相同”，但是对于无限集来说，元素的“个数”指的是什么呢? 称一个集合 X 为有限(finite)集，当且仅当它和某个 $n\in\mathbb{N}$ 有相同的势（这里的 n 代表集合$\{0, 1, \cdots, n-1\}$，它有 n 个元素）. 否则，就称 X 是无限集. 例如，$\mathbb{N}$是无限集. 称 X 为可数(countable)集，当且仅当存在从$\mathbb{N}$映上到 X 的函数 f，如果 X 还是无限的，则称 X 是可数无限(countably infinite)集. 一个集合是不可数(uncountable)集，当且仅当它不是可数的.

例如，$\mathbb{N}$与$\mathbb{N}\times\mathbb{N}$有相同的势，因为存在函数 $f(m, n):=2^m(2n+1)-1$，它是一个从$\mathbb{N}\times\mathbb{N}$映射到$\mathbb{N}$的一一对应的函数.

称 X 比 Y 的势小(smaller cardinality)，当且仅当存在一个从 X 映射到 Y 的 1－1 的函数，但是不存在从 X 映上到 Y 的函数. 下面的事实说明这样的定义是一致的.

1.4.1 等价定理 设 A，B 是两个集合，f 是从 A 映射到 B 的 1－1 的函数，g 是从 B 映射到 A 的 1－1