现代数学基础丛书 159

算法数论

（第二版）

裴定一　祝跃飞　编著

科学出版社

现代数学基础丛书 159

算法数论

(第二版)

裴定一　祝跃飞　编著

科学出版社

北京

内 容 简 介

本书论述了算法数论的基本内容，其中涉及同余式、二次剩余、特征、连分数、代数数域、椭圆曲线、素性检验、大整数因子分解算法、椭圆曲线上的离散对象、超椭圆曲线、格理论等分支，也介绍了这些知识在密码学中的一些应用. 本书的特点是内容涉及面广，在有限的篇幅内，包含了必要的预备知识和数学证明，尽可能形成一个比较完整的体系. 本书的部分内容曾多次在中国科学院研究生院信息安全国家重点实验室和广州大学作为硕士研究生教材使用.

本书可作为信息安全、数论等专业的研究生教材，以及相关专业的研究人员、高等学校的教师和高年级学生的参考书.

图书在版编目(CIP)数据

算法数论/裴定一，祝跃飞编著. —2 版. —北京：科学出版社，2015.7
（现代数学基础丛书；159）

ISBN 978-7-03-045332-7

Ⅰ. ①算… Ⅱ. ①裴… ②祝… Ⅲ. ①算法理论 Ⅳ. ①O241

中国版本图书馆 CIP 数据核字(2015) 第 186575 号

责任编辑：李静科 / 责任校对：张凤琴
责任印制：徐晓晨 / 封面设计：陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京教图印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*
2002 年 9 月第 一 版 开本：720 × 1000 1/16
2015 年 9 月第 二 版 印张：15 1/2
2015 年 9 月第一次印刷 字数：293 000

定价：78.00 元

(如有印装质量问题，我社负责调换)

《现代数学基础丛书》编委会

《现代数学基础丛书》序

对于数学研究与培养青年数学人才而言，书籍与期刊起着特殊重要的作用．许多成就卓越的数学家在青年时代都曾钻研或参考过一些优秀书籍，从中汲取营养，获得教益．

20世纪70年代后期，我国的数学研究与数学书刊的出版由于文化大革命的浩劫已经破坏与中断了 10 余年，而在这期间国际上数学研究却在迅猛地发展着．1978 年以后，我国青年学子重新获得了学习、钻研与深造的机会．当时他们的参考书籍大多还是 50 年代甚至更早期的著述．据此，科学出版社陆续推出了多套数学丛书，其中《纯粹数学与应用数学专著》丛书与《现代数学基础丛书》更为突出，前者出版约 40 卷，后者则逾 80 卷．它们质量甚高，影响颇大，对我国数学研究、交流与人才培养发挥了显著效用．

《现代数学基础丛书》的宗旨是面向大学数学专业的高年级学生、研究生以及青年学者，针对一些重要的数学领域与研究方向，作较系统的介绍．既注意该领域的基础知识，又反映其新发展，力求深入浅出，简明扼要，注重创新．

近年来，数学在各门科学、高新技术、经济、管理等方面取得了更加广泛与深入的应用，还形成了一些交叉学科．我们希望这套丛书的内容由基础数学拓展到应用数学、计算数学以及数学交叉学科的各个领域．

这套丛书得到了许多数学家长期的大力支持，编辑人员也为其付出了艰辛的劳动．它获得了广大读者的喜爱．我们诚挚地希望大家更加关心与支持它的发展，使它越办越好，为我国数学研究与教育水平的进一步提高做出贡献．

杨　乐

2003 年 8 月

第二版前言

本书的第一版于 2002 年 9 月出版, 目的是介绍与公钥密码相关的数论算法. 近年来, 公钥密码有了很多新的发展, 本版新增以下内容.

1. 二次剩余和格理论是数论中两个古老的分支, 近年来在密码理论中得到了重要应用. 第二版在第一版基础上增添了这两方面的内容. 3.4 节介绍二次剩余假设的概念. 基于二次剩余假设这一数学难题, 8.5 节构造了一个概率公钥密码, 并证明它具有多项式安全.

2. 格密码是密码学界研究的热点问题之一, 第 13 章 "格" 是第一版附录中的 A.5 节的扩充. 本章介绍格的基本理论及其在密码学中的应用, 包括格的基本概念和 LLL 算法, 以及 LLL 算法在背包问题求解和小指数 RSA 密码算法分析中的应用, 最后介绍两类基于格中数学难题设计的公钥密码体制, 包括 NTRU 密码体制和基于 LWE 难题的全同态加密体制.

此外, 第二版中还增加了 "名词索引".

作　者

2015 年 4 月

第一版前言

算法数论是一门对数论问题进行算法设计和算法分析的学科. 它的历史可以追溯到古希腊 Eratosthenes 氏筛法构造素数表. 但它真正成为一门学科, 是在 20 世纪中叶, 一方面是由于计算机科学的发展和计算复杂度理论的建立为算法数论奠定了理论基础; 另一方面是数论发展的内部推动力, 如对数论中某些问题 (如一些猜想) 给出肯定与否的回答的过程中, 收集依据时涉及一些大数据量的实例的验证, 而这已经超出了人们的手算能力, 只能借助计算机编程来完成; 更重要的是由于一些基于数论的公钥密码方案的提出和对其攻击所涉及的一些数论问题求解算法的发现.

公钥密码是在 20 世纪 70 年代中期提出的一类新型的密码, 它尤其适合在计算机网络环境下使用, 具有加密信息、管理密钥和数字签名等功能, 能保证信息的机密性、完整性和不可否认性. 迄今为止所提出的公钥密码, 其安全性都建立在某个数学难题的基础之上, 所谓 "数学难题", 确切地说是求解这个数学问题, 目前还没有多项式时间的算法被发现. 例如, 大整数因子分解、有限域或椭圆曲线离散对数等问题. 只要选择适当的参数, 在现有的技术条件下, 这些问题都是很难解决的, 这就为相应的公钥密码的安全性奠定了基础. 在解决这些难题方面所取得的任何重大进展, 都会对相应的公钥密码的使用产生巨大的影响.

RSA 公钥密码、ElGamal 公钥密码和椭圆曲线公钥密码是目前影响最大的三类公钥密码. 前者是在 20 世纪 70 年代中叶提出来的, 它的安全性依赖于大整数因子分解的难度, 后两者的安全性分别依赖于计算有限域离散对数和椭圆曲线离散对数的难度. 椭圆曲线公钥密码是 20 世纪 80 年代中叶提出来的, 由于其自身具有一些其他公钥体制无法比拟的优势, 近十年来已成为公钥密码研究的一个十分活跃的方向, 研究所获得的许多有关的椭圆曲线的算法, 大大丰富了算法数论的理论.

因子分解和离散对数是算法数论研究的两个核心问题. 本书的主要内容是介绍这两个问题的基本理论, 以及迄今为止所提出的主要算法的基本原理. 这部分内容包含在第 9~11 章. 第 9 章介绍 Miller-Rabin 概率型素性检验方法, 以及分别利用特征和、椭圆曲线的确定型检验方法. 第 10 章重点介绍椭圆曲线因子分解方法及数域筛法. 第 11 章包含有关有限域及椭圆曲线上的离散对数的主要结果. 其余各章 (除第 8 章) 都是为这最后三章作准备的. 第 1~5 章介绍初等数论的有关知识, 第 6 章介绍代数数论的有关预备知识, 第 7 章介绍椭圆曲线的有关预备知识. 第 8 章介绍前五章的初等数论知识在密码学中的一些应用. 为了本书的系统性, 添加一个附录, 介绍一些代数和有限域的一些算法.

本书的选材是经过精心考虑的, 内容涉及面很广, 但在有限的篇幅内, 包含了必要的预备知识和数学证明, 尽可能形成一个较完整的体系. 考虑到信息安全专业的研究生有来自数学本科和非数学本科两类, 在利用本书时, 可以根据需要, 选择不同的章节组成一个学期的教学. 对于来自数学本科的学生, 前五章可以较快地通过, 而把重点放在后面几章. 对于来自非数学本科的学生, 第 7~11 章有关代数数论和椭圆曲线的章节可以考虑不讲. 本书的部分内容曾多次在中国科学院研究生院信息安全国家重点实验室和广州大学作为硕士研究生教材.

本书的编写和出版得到国家自然科学基金跨学科重点项目 “电子商务系统中的信息安全理论和技术的研究” (批准号 19931010), 国家 973 项目 “信息与网络安全体系结构” (批准号 G1999035804) 和 “中国科学院研究生教材基金” 的资助, 特此感谢.

作　者

2001 年 4 月

目　录

第 1 章　整数的因子分解

1.1　唯一分解定理

数论是研究自然数 $1,2,3,\cdots$ 性质的一门数学分支. 自然数是人们日常生活中用得最多的一类数. 历史上, 人们很早就开始研究数论, 它已成为内容十分丰富的一个分支. 数论在信息安全、计算机科学、数字信号处理等现代科技领域有重要的应用, 所以, 数论至今仍是一门充满活力、蓬勃发展的分支.

通常, 用 $\mathbb{Z}$ 表示整数集合, 整数即为

$$0, \quad \pm 1, \quad \pm 2, \cdots .$$

自然数就是正整数.

定理 1.1　*设 a 和 b 为整数, $b>0$, 则存在整数 q 和 r, 使*

$$a = qb + r, \quad 0 \leqslant r < b,$$

r 称为 b 除 a 所得的最小正剩余.

证明　以 $\left[\dfrac{a}{b}\right]$ 表示不超过分数 $\dfrac{a}{b}$ 的最大整数, 则

$$0 \leqslant a - \left[\frac{a}{b}\right] b < b,$$

取 $q = \left[\dfrac{a}{b}\right], r = a - \left[\dfrac{a}{b}\right] b$, 即证得定理.

当 b 除 a 的最小正剩余 r 为零时, 称 b 为 a 的因子, a 为 b 的倍数, 记为 $b|a$.

若 b 为 a 的因子, $b \neq 1, b \neq a$, 这时称 b 为 a 的真因子, 显然有 $0 < |b| < |a|$, 这里 $|a|$ 为 a 的绝对值.

若 $b \neq 0$, $c \neq 0$, 显然有:

(1) 若 $b|a$, $c|b$, 则 $c|a$;

(2) 若 $b|a$, 则 $bc|ac$;

(3) 若 $c|d$, $c|e$, 则对任意 m, n 有 $c|dm + en$.

自然数 $p(\neq 1)$, 若仅以 1 和自身 p 为其因子, 称 p 为素数. 非素数的自然数 $n\ (\neq 1)$ 称为复合数.

设 M 为整数的一个子集合, 如果它对加、减法封闭, 即若 $m, n \in M$, 则 $m \pm n \in M$, 这时称 M 为模. a 为任一整数, a 的所有的倍数就组成一个模. 相反的结论也成立, 即如下定理.

定理 1.2　*任一非零模, 必为一正整数的诸倍数组成的集合.*

证明　设 d 为该模中最小正整数, 则模中其他数必为 d 之倍数. 若不然, 设 n 为模中 d 之非倍数, 由定理 1.1, 存在整数 q 及 r, 使

$$n = qd + r, \quad 0 < r < d.$$

由于 $r = n - qd$ 也属于此模, 这与 d 为该模中最小正整数的假设相矛盾, 故模中其他各数都为 d 的倍数. 因为 d 在模中, 所以 d 的任一倍数也在模中. 定理即证.

命 a, b 为二整数, 集合

$$\{ma + nb \mid m, n \in \mathbb{Z}\}$$

即为一模, 此模中最小正整数 d 称为 a, b 的最大公因子, 记为 $d = (a, b)$.

由定理 1.2 的证明, 不难证得下述定理.

定理 1.3　*(a, b) 具有下述性质:*

(1) *有整数 x, y, 使 $(a, b) = ax + by$;*

(2) *对任二整数 x, y, 必有 $(a, b) | ax + by$;*

(3) *若 $c|a$, $c|b$, 则 $c|(a, b)$.*

由于 (3), 也称 (a, b) 为 a, b 的最大公因子.

定理 1.4　*设 p 为素数且 $p|ab$, 则 $p|a$ 或 $p|b$.*

证明　若 $p \nmid a$, 则 $(a, p) = 1$, 由定理 1.3 知有二整数 x, y, 使

$$ax + py = 1,$$

所以

$$abx + pyb = b.$$

由 $p|ab$ 可知 $p|b$, 证毕.

定理 1.5(唯一分解定理)　*任一自然数 n 皆可唯一地表为素数之积*

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}. \tag{1.1}$$

这里, $p_1 < p_2 < \cdots < p_k$ 为素数, $a_1, a_2, \cdots, a_k$ 为自然数.

证明　首先证明 n 可以表为素数之积, 然后再证明上述表法唯一.

若 n 为素数, 定理显然成立. 当 n 不是素数时, 设 p_1 是 n 的最小的真因子, 则 p_1 一定是素数, 因 p_1 的真因子也是 n 的真因子, 所以 p_1 不能有真因子. 设

$n = p_1 n_1$ $(1 < n_1 < n)$, 对 n_1 重复上述推理, 得 $n = p_1 p_2 n_2$, p_2 为素数, $1 < n_2 < n_1$, 继续执行此法, 得 $n > n_1 > n_2 > \cdots > 1$, 此做法最多不能超过 n 次, 最后必得

$$n = p_1 p_2 \cdots p_l,$$

也可排为式 (1.1) 中的形式.

今设

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k} = q_1^{b_1} q_2^{b_2} \cdots q_l^{b_l}$$

为 n 的两个分解式, 其中 $p_1 < p_2 < \cdots < p_k$, $q_1 < q_2 < \cdots < q_l$ 都为素数, 利用定理 1.4, 任一 p_i 必为某一 q_j, 任一 q_i 也必为某一 p_j, 故 $k = l, p_i = q_i$ $(1 \leqslant i \leqslant k)$, 又若 $a_1 > b_1$, 则

$$p_1^{a_1 - b_1} p_2^{a_2} \cdots p_k^{a_k} = p_2^{b_2} \cdots p_k^{b_k},$$

左边为 p_1 的倍数, 右边不是 p_1 的倍数, 这是不可能的, 同样 $a_1 < b_1$ 也不可能, 故 $a_1 = b_1$. 类似地, 可证得 $a_i = b_i$ $(i = 1, 2, \cdots, k)$, 唯一性得证.

给定一自然数 n, 当它很大时, 例如, 一百多位的十进制数, 要将它因子分解, 实非易事. 在第 10 章将讨论一些大整数因子分解的算法, 随之而来的一个问题是如何判断一个数是否是素数, 在第 9 章将讨论几个素性判断的方法.

1.2 辗转相除法 (欧氏除法)

若 a, b 为二自然数, $a \geqslant b$, 以 (a, b) 表示 a 和 b 的最大公因子. 由定理 1.3 知, 有二整数 x, y, 使

$$(a, b) = ax + by.$$

如何计算 (a, b), 又如何找到上述 x 和 y, 定理 1.1 实际上已经给出了所要的算法.

首先用 b 除 a 得到商 q_0, 余数 r_0, 即

$$a = q_0 b + r_0, \quad 0 \leqslant r_0 < b. \tag{1.2}$$

如果 $r_0 = 0$, 那么 b 是 a 的因子, a, b 的最大公因子就是 b. 如果 $r_0 \neq 0$, 用 r_0 除 b 得到商 q_1, 余数 r_1, 即

$$b = q_1 r_0 + r_1, \quad 0 \leqslant r_1 < r_0. \tag{1.3}$$

如果 $r_1 = 0$, 那么 r_0 除尽 b, 由式 (1.2) 知, r_0 也除尽 a, r_0 是 a, b 的公因子. 反之, 任何一个除尽 a, b 的数, 由式 (1.2) 知, 也除尽 r_0, 因此 r_0 是 a, b 的最大公因子. 如果 $r_1 \neq 0$, 则用 r_1 除 r_0 得到商 q_2, 余数 r_2, 即

$$r_0 = q_2 r_1 + r_2, \quad 0 \leqslant r_2 < r_1. \tag{1.4}$$

如果 $r_2 = 0$, 那么由式 (1.3) 可知, r_1 是 r_0, b 的公因子, 由式 (1.2) 知, r_1 也是 a, b 的公因子. 反之, 如果一整数除得尽 a, b, 那么由式 (1.2) 知, 它一定除得尽 r_0 , 由式 (1.3) 知, 它一定除得尽 r_1, 所以 r_1 是 a, b 的最大公因子.

若 $r_2 \neq 0$, 再用 r_2 除 r_1, 重复上述过程, 依次得到 $b > r_0 > r_1 > r_2 > \cdots$, 逐步小下来, 而又都非负. 经过有限步后, 一定会有某个 r 为零. 若设 r_n 是第一个出现的零, 则 r_{n-1} 就是 a, b 的最大公因子. 所得到的一串算式为

$$\begin{aligned}
a &= q_0 b + r_0, \\
b &= q_1 r_0 + r_1, \\
r_0 &= q_2 r_1 + r_2, \\
r_1 &= q_3 r_2 + r_3, \\
&\cdots\cdots \\
r_{n-3} &= q_{n-1} r_{n-2} + r_{n-1}, \\
r_{n-2} &= q_n r_{n-1}.
\end{aligned}$$

由第一式可得

$$r_0 = a - q_0 b,$$

由第二式可得

$$r_1 = b - q_1 r_0 = -q_1 a + (1 + q_0 q_1) b,$$

一般地, 对任一 r_i $(0 \leqslant i \leqslant n-1)$, 都有二整数 x_i, y_i, 使

$$r_i = x_i a + y_i b.$$

由于

$$\begin{aligned}
r_i &= r_{i-2} - q_i r_{i-1} \\
&= (x_{i-2} a + y_{i-2} b) - q_i (x_{i-1} a + y_{i-1} b) \\
&= (x_{i-2} - q_i x_{i-1}) a + (y_{i-2} - q_i y_{i-1}) b,
\end{aligned}$$

所以有递推公式

$$\begin{aligned}
&x_0 = 1, \qquad x_1 = -q_1, \qquad x_i = x_{i-2} - q_i x_{i-1}, \\
&y_0 = -q_0, \qquad y_1 = 1 + q_0 q_1, \qquad y_i = y_{i-2} - q_i y_{i-1}.
\end{aligned}$$

这样, 可以找到二整数 x, y, 使

$$(a, b) = ax + by.$$

看一个例子: 求 4862 和 2156 的最大公因子, 则有

$$
\begin{aligned}
4682 &= 2 \times 2156 + 550,\\
2156 &= 3 \times 550 + 506,\\
550 &= 506 + 44,\\
506 &= 11 \times 44 + 22,\\
44 &= 2 \times 22.
\end{aligned}
$$

可见 $(4862, 2156) = 22$, 利用上述算式可得

$$
\begin{aligned}
550 &= 4862 - 2 \times 2156,\\
506 &= -3 \times 4862 + 7 \times 2156,\\
44 &= 4 \times 4862 - 9 \times 2156,\\
22 &= -47 \times 4862 + 106 \times 2156.
\end{aligned}
$$

称上述求 a, b 的最大公因子的算法为辗转相除法, 或欧几里得除法.

考虑辗转相除法所需的比特计算量. 仍设 $a \geqslant b$, 若 a 和 b 用二进制表示的长度分别为 k 和 l, 则 $k \leqslant \log_2 a + 1$, $l \leqslant \log_2 b + 1$. 用 b 除 a 得到商和余数, 这个带余除法所需的比特计算量为 $O(kl)$(这里 $O(kl)$ 表示一个 $\leqslant c \cdot kl$ 的量, 其中 c 为一个不依赖 k 和 l 的常数), 也可表为 $O(\lg^2 a)$. 还需要知道带余除法要做多少次.

我们有 $r_{j+2} < \dfrac{1}{2} r_j$.

首先来证明这个论断. 若 $r_{j+1} \leqslant \dfrac{1}{2} r_j$, 则 $r_{j+2} < r_{j+1} \leqslant \dfrac{1}{2} r_j$, 即证. 若 $r_{j+1} > \dfrac{1}{2} r_j$, 则 $r_j = r_{j+1} + r_{j+2}$, 同样有 $r_{j+2} < \dfrac{1}{2} r_j$.

以上论断表示, 做两次带余除法可将余数缩小一半. 要得到 (a, b), 所要做的带余除法的次数不会超过 $2[\log_2 a] = O(\lg a)$, 因而辗转相除法所需的比特计算量为

$$
O(\lg^2 a) \times O(\lg a) = O(\lg^3 a).
$$

给定自然数 a, b, 若已知它们的因子分解

$$
\begin{aligned}
a &= p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}, \quad \alpha_i \geqslant 0 \quad (1 \leqslant i \leqslant s);\\
b &= p_1^{\beta_1} p_2^{\beta_2} \cdots p_s^{\beta_s}, \quad \beta_i \geqslant 0 \quad (1 \leqslant i \leqslant s),
\end{aligned}
$$

则

$$
(a, b) = p_1^{\min(\alpha_1, \beta_1)} p_2^{\min(\alpha_2, \beta_2)} \cdots p_s^{\min(\alpha_s, \beta_s)}.
$$

若以 $[a, b]$ 表示 a, b 的最小公倍数 (即 $[a, b]$ 是 a 和 b 的倍数, 且任一 a 和 b 的公倍数都是 $[a, b]$ 的倍数), 则

$$
[a, b] = p_1^{\max(\alpha_1, \beta_1)} p_2^{\max(\alpha_2, \beta_2)} \cdots p_s^{\max(\alpha_s, \beta_s)}.
$$