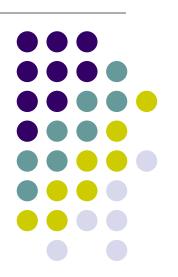
# 集合计数

离散数学: 第八讲



#### 提要

NA AND THE STATE OF THE STATE O

- 集合的大小
- 无限集合
- 等势与优势关系
- 集合计数
- 容斥原理

#### 自然数与无穷集合



God made the integers; all else is the

work of man.

Leopold Kronecker

## 回顾: 无穷公理



- 无穷公理(Axiom of Infinity, 在ZF系统内):  $\exists A \big( \emptyset \in A \land (\forall x \in A)(x^+ \in A) \big)$
- 以往按照von Neumann的定义, $0 = \emptyset$ ,  $n + 1 = n^+$ ,从而可以定义出单个的自然数,但不能说明全体自然数集合 $\mathbb{N}$ 的存在性,而由无穷公理可以定义 $\mathbb{N}$
- 在集合论中: № ≝ ∩ {A|A为归纳集}

#### 回顾: 从集合构造自然数



■ 设x为集合,x的后继(successor)x+指x  $\cup$  {x},

是von Neumann的定义

■ 设A为集合,称A为归纳集(inductive set)指:

$$\emptyset \in A \land (\forall x \in A)(x^+ \in A)$$

## 自然数的Peano公理系统



- Peano系统的公理(i.e. 自然数五公设)为:
  - $\circ$  **Ax.1 0**  $\in$  N
  - $\bigcirc \quad \mathbf{Ax.2} \quad \boldsymbol{n} \in \mathbb{N} \to \boldsymbol{n}_{\mathrm{Def}}^{\mathrm{The \ sign}} \text{is read } \text{minus,} < \text{is read } \text{is less than, and} >$

  - $\circ \quad \mathbf{Ax.5} \quad \mathbf{0} \in A \land (\forall x \in A)(x^+ \in A) \rightarrow (\forall x \in \mathbb{N})(x \in A)$
- 由此可由集合构造出Peano算术(PA): (N, 0, +)





- 对于自然数的von Neumann定义,可定义:
  - $m \le n \stackrel{\text{def}}{=} m \subseteq n$ , 于是有以下命题成立:
  - (1)  $n + 1 = \{0, 1, 2, \dots, n\}$  is read minus, < is read is less than, and > $\neg : b - a = \mathbb{N}[x \in ](x + a = b).$
  - (2)  $n \in n + 1$
  - (3)  $n \leq n$   $q \in N$
  - (4)  $n \le m \le 1 \rightarrow n \le 1$ ;  $n \le m \le n \rightarrow n = m$  $\supset_x \cdot x + 1 \in k :: \supset \cdot N \supset k$
  - (5)  $m \leq n \vee n \leq m$

#### 我们怎么比较集合的大小



• "数得清"的我们就数元素个数。

- "无数"的怎么办?
  - "常识"不一定经得起追问。
- 集合的大小称为集合的"势" (cardinality)
  - 集合S的势记为|S|





#### 有限与无限:"宇宙旅馆"



啊?客满啦?

没关系,我让现在住在 k 号房间的客人移到 k+1号。你就住进第1号房间吧!

#### 集合的等势关系



- 等势关系的定义:
  - 如果存在从集合A到集合B的双射,则称集合A与B等势。
  - 集合A与B等势记为: A≈B, 否则A≈B
  - A≈B意味着: A, B中的元素可以"一一对应"。
  - 要证明A≈B,找出任意一个从A到B的双射即可。
- "等势"的集合就被认为是"一样大"

#### 自然数集是无限集



■ 证明: 自然数集 № 是无穷集

反设N有穷,从而存在n以及双射函数 $f:n \to N$ ,令

 $m = f(0) + f(1) + \cdots + f(n-1) + 1$ , 从而有 $\forall x \in$ 

 $n, f(x) \neq m$ ,故f非满射,矛盾!故N为无穷集.  $\square$ 

#### 可列集



- 上述定义中,可列集的直观概念可以看作集合的元素可以按确定的顺序线性排列,所谓"确定的"顺序是指对序列中任一元素,可以说出它"前一个"、"后一个"元素是什么
- 例如:整数集与自然数集等势,故ℤ为可列集 0,-1,1,-2,2,-3,3,-4,···

#### 可列集



证明:构造如下f: Z → N,易见f为双射.

#### 将 Z 中元素以下列顺序排列并与 N 中元素对应:

$$Z: 0 -1 1-2 2 -3 3 ...$$
 $\downarrow \qquad \downarrow \qquad N: 0 1 2 3 4 5 6 ...$ 

则这种对应所表示的函数是:

$$f: Z \to N, f(x) = \begin{cases} 2x & \ge 0 \\ -2x - 1 & x < 0 \end{cases}$$

#### 有限集与无限集



- S是有限集合, iff. 存在自然数n, 使得S与n等势
- S不是有限集合(即:无限集),iff.存在S的真子集S',使得S与S'等势
  - ⇒ S一定包含一个与自然数集合等势的子集 $M = \{a_1, a_2, a_3, ...\}$  (这实际上意味着:自然数集是"最小的"无限集)

令 $S'=S-\{a_1\}$ ,可以定义 $f:S\to S'$ 如下:

对于任意 $x \in M$ ,  $f(a_i) = a_{i+1}$ ; 对于任意 $x \in S-M$ , f(x) = x

显然这是双射,即S与其真子集S等势

 $\leftarrow$  假设S是有限集,令|S|=n,则给S任意的真子集S',若|S'|=m,必有m<n,因此从S '到S的任一单射不可能是满射。

#### 有穷与无穷:差别不仅是数量



- 等势的涵义是两个集合元素的个数"一样多"
- 整体一定大于部分吗?
- Galileo佯谬(Galileo's paradox, 1638):
  - (1) 令 $\mathbb{N}^{(2)} = \{0^2, 1^2, 2^2, 3^2, \cdots\}$ ,显然: $\mathbb{N}^{(2)} \subset \mathbb{N}$ ;但G. Galileo发现 $\mathbb{N}^{(2)}$ 与 $\mathbb{N}$ 中的元素——对应: $\mathbb{P}^{(2)}$ 令 $\mathbb{N}^{(2)}$ 如下: $\mathbb{N}^{(2)}$ 如下: $\mathbb{N}^{(2)}$ 如下: $\mathbb{N}^{(2)}$ 如下: $\mathbb{N}^{(2)}$
  - (2) 令 $\mathbb{N}^* = \{0,1^1,2^2,3^3,4^4,5^5,\cdots\}$ , 易见 $\mathbb{N} \approx \mathbb{N}^*$

## 有穷与无穷:差别不仅是数量



■ Hilbert 佯谬: {0, 1, 2, ···} ≈ {1, 2, 3, ···}



客满了?没关系, 让现在住在 k 号房 间的客人移到 k+1 号。新来的客人就 住进第1号房间吧!

#### 与自然数有关的若干命题



- (1) 自然数*n*的任何真子集为有限集
- (2) 任何自然数不等势于其真子集
- (3) 若集合A有穷,则A不与其任何真子集等势
- (4) 若集合A与其某个真子集等势,则A无穷
- 其中(2)即为"鸽笼原理"





■ 对于互异的 $a,b \in \mathbb{R}$ 和互异的 $c,d \in \mathbb{R}$ ,有:

 $[a,b] \approx [c,d], (a,b) \approx (c,d)$ 



对任何  $a, b \in R$ , a < b, [0,1] ≈ [a,b].

双射函数  $f: [0,1] \rightarrow [a,b], f(x) = (b-a)x+a$ 

类似地可以证明, 对任何  $a, b \in R$ , a < b, f(0,1)≈(a,b).

 $\tau(X)$ 

#### 证明无限集等势的例子

 $\tau(X)$ 

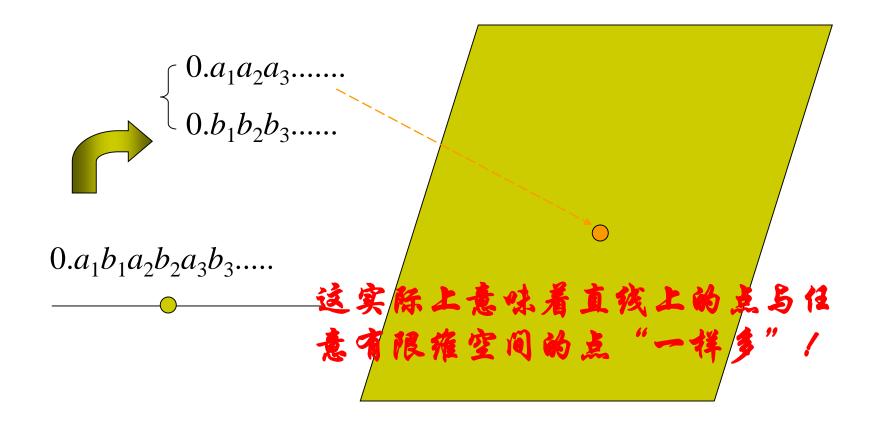


■ 命题(Riemann): 设 $a \neq b$ , 则(a, b)  $\approx \mathbb{R}$ 

 $(0,1)\approx R$ . 其中实数区间  $(0,1)=\{x|x\in R\land 0< x<1\}$ . 令 双射函数  $f:(0,1)\to R$ ,  $f(x)=\tan\pi\frac{2x-1}{2}$ 



## 直线上的点集与平面上的点集等势



#### 集合的优势关系



- 如果存在从集合A到集合B的单射,则称"集合B优势于集合A"
- 集合B优势于集合A 记为 A≤•B
- 如果集合B优势于集合A,且B与A*不等势*,则 称 "集合B真优势于集合A",记为A≺•B
- 实数集合真优势于自然数集
- 例子:对任意集合A,A的幂集真优势于集合A

#### 集合优势关系的性质



- 自反性: 恒等函数
- 若A≤•B, 且B≤•A, 则A≈B (比较:反对称性)
   (Cantor-Bernstein定理)
- 传递性: 单射的复合仍然是单射





- 有时候找双射不太容易
  - 证明实数集的两个子集(0,1)和[0,1]等势。

关键是如何安排在[0,1]中但不在(0,1)中的0和1。

想象那个"宇宙旅馆"。我们可以取(0,1)的一个与自然数集合等势的子集 $(-定有)\{a_1,a_2,a_3,...\}$ ,"腾出"前两个位置安排(0,1)0和1

一种证法: 
$$f(x) = \begin{cases} \frac{1}{2} & x = 0\\ \frac{1}{2^2} & x = 1\\ \frac{1}{2^{n+2}} & x = \frac{1}{2^n}, n = 1, 2, 3...\\ x & x \neq 1 \end{cases}$$

# 优势关系的反对称性用于证明等势(续)

- 证明实数集的两个子集(0,1)和[0,1]等势。
- 分别找两个一对一的映射往往比找一个双射容易

$$f:(0,1) \to [0,1]: f(x) = x$$

$$g:[0,1] \to (0,1): g(x) = \frac{1}{2} + \frac{x}{4} \quad 注意: g([0,1]) = [\frac{1}{2}, \frac{3}{4}]$$





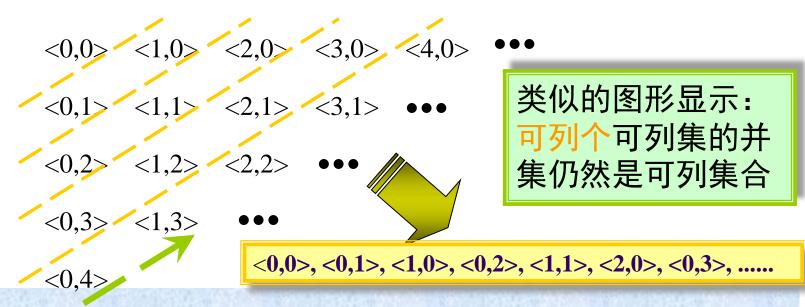
#### 集合A的基数

- 若A与自然数n等势,则card A = n
- 若A与自然数集合N等势,则card  $A = \aleph_0$
- 若A与实数集合R等势,则card A = ※
- 如果存在从A到N的单射,则称A为可数集,或可 列集。[card A ≤ ℵ₀]





所有的整数对构成的集合与自然数集等势



$$l(m,n) = \sum_{i=1}^{m+n} i + (m+1)$$

#### 实数集不是可列集



- 命题: 实数集非可列集
- 证明(Cantor's diagonalization argument, 1891):

由于 $\mathbb{R} \approx [0,1]$ ,故只需要说明[0,1]之间的实数点集不可列即可。首先约定实数 $x \in [0,1]$ ,令 $x = 0.x_1x_2x_3\cdots(0 \le x_i \le 9)$ ,对于无限循环小数 $0.249999\cdots$ 与 $0.250000\cdots$ 统一只采用后者的表示法表示;



■ 证明(Cantor's diagonalization argument)(续):

假设[0,1]之间用上述方法表示的实数 可列,则[0,1]上

的值可列举为:

 $0.\mathbf{b_{11}}\mathbf{b_{12}}\mathbf{b_{13}}\mathbf{b_{14}}\dots$ 

 $0.b_{21}b_{22}b_{23}b_{24}...$ 

 $0.b_{31}b_{32}b_{33}b_{34}...$ 

 $0.b_{41}b_{42}b_{43}b_{44}...$ 

፥

今取实数 $y \in [0,1]$ ,将其表为 $0.b_1b_2b_3$ …,并令 $b_i \neq b_{ii}$  ( $i = 1,2,3,\dots$ )。易见,y与上表中任何一个值均不等,上述假设错误。即实数集ℝ是不可列集.□

#### 例子



#### 一个例子:

$$r_1 = 0.5105110...$$

$$r_2 = 0.4132043...$$

$$r_3 = 0.8245026...$$

$$r_4 = 0.2330126...$$

$$r_5 = 0.4107246...$$

$$r_6 = 0.9937838...$$

$$r_7 = 0.0105135...$$

$$r_1 = 0.5105110...$$

$$r_2 = 0.4 132043...$$

$$r_3 = 0.82 45026...$$

$$r_4 = 0.233 \, \mathbf{0} \, 126 \dots$$

$$r_5 = 0.4107246...$$

$$r_6 = 0.9937838...$$

$$r_7 = 0.0105135 \dots$$

#### 幂集的基数



■ 命题:  $\mathcal{P}(A) \approx \{0,1\}^A = \{f|f:A \to \{0,1\}\}$ 

故 $\mathscr{P}(A) \sim \{0,1\}^A i.e. \mathscr{P}(A) \sim A_I \quad \square$ 

#### 康托尔定理



- Cantor 定理(1891):
  - **(1)** N ≈ R
  - (2) 对于任意集合A,  $A \approx P(A)$
- 证明: (1) 参见对角线法;
- (2) 证明非 $A \sim \mathcal{P}(A)$ ,

反设 $f: A \xrightarrow{1-1} \mathscr{P}(A)$  令  $B = \{x \in A \mid x \notin f(x)\}$ ,显然 $B \in \mathscr{P}(A)$ ,但 $B \notin Ran(f)$ ,这是因为若B = f(a)则 $a \in B \leftrightarrow a \notin f(a) \leftrightarrow a \notin B$ 矛盾!故f非onto矛盾。

# 0到1之间有理数多还是无理数多



南京大学小百合站 -- 主题文章阅读 [讨论区: Pictures][回帖预定]

#### 添加标签

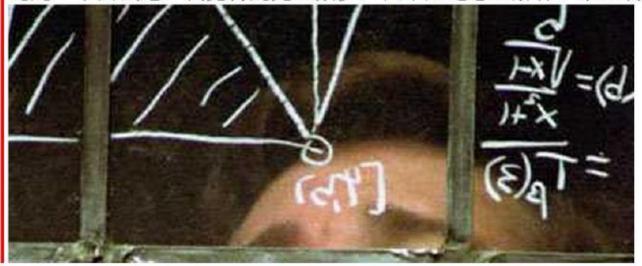
追踪此人

[本篇全文] [回复本文] [本篇作者: xuq(男生)] [本篇人气: 2251]

发信人: xuq (mossad), 信区: Pictures 标 题: 0到1之间有理数多还是无理数多

发信站: 南京大学小百合站 (Tue Mar 15 16:50:37 2011)

这是一个面试题,我觉得是无理数多,但不知道怎么解释。求证明!





#### ■ 证明:

(1) 易构造如下双射  $f: \mathbb{Q} \to \mathbb{N}$ ,故  $\mathbb{Q} \approx \mathbb{N}$ ,为可列集



#### ■ 证明(续):

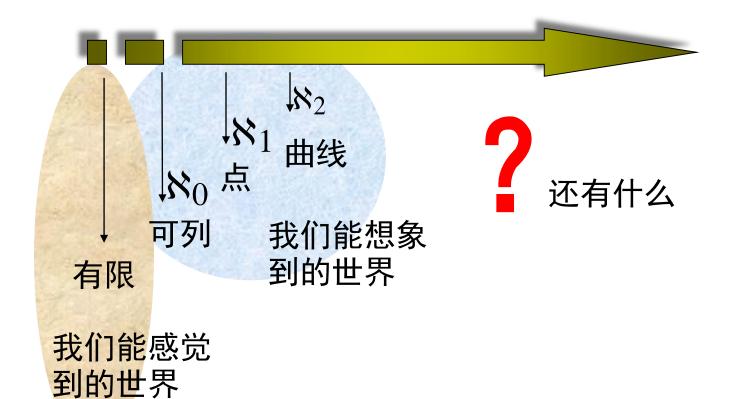
(2) 用Cantor对角线法易证ℝ – ℚ不可列,故

$$|\mathbb{R} - \mathbb{Q}| = \aleph$$

因此 $\mathbb{Q} < \mathbb{R} - \mathbb{Q}$ ,即[0,1]内无理数的基数比有理数大.

# 集合的"大小" - 基数





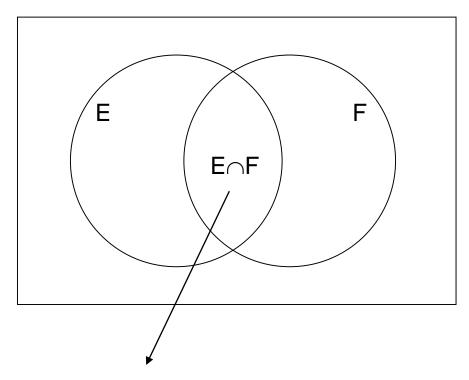
#### 重要的等势优势关系



- $\blacksquare \mathbb{N} \approx \mathbb{Z} \approx \mathbb{Q} \approx \mathbb{N} \times \cdots \times \mathbb{N}$
- $\mathbb{R} \approx [a,b] \approx (c,d) \approx \{0,1\}^{\mathbb{N}} \approx P(\mathbb{N}) \approx (0,1)$
- N < · R</p>
- $\blacksquare A \prec P(A) \prec P(A) \prec \cdots$

### 集合用于分类





既学英语, 又学法语的同学

将属于某个集合的元素理解为"具有某种性质"的对象,则属于该集合的补集的元素则是"不具有某种性质"的对象。

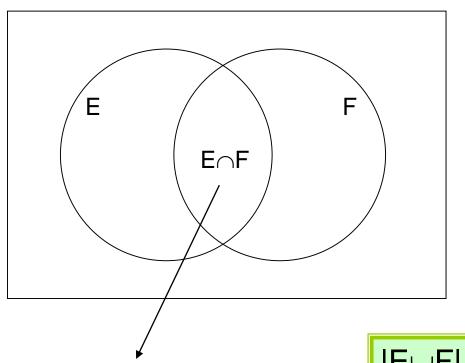
#### 例如:

将全班同学的集合视为全集。

其子集E是学英语的同学的集合,或理解为满足性质E的对象的集合。类似地,F是学法语的同学的集合,即满足性质F的对象的集合。







假设全班共100人,记为

|U| = 100

学英语的50人,学法语的30人,分别记为:

|E| = 50; |F| = 30

显然,只要 $E \cap F \neq \emptyset$ ,既不 学英语,也不学法语的人 数并非20人。

既学英语, 又学法语的同学

 $|E \cup F| = (|E| + |F|) - |E \cap F|$ 

# 多少种排法?



- 将0,1,2,...,9排成一列,要求第1个数字大于1,最后一个数字小于8,共有多少种排法?
  - 这10个数字所有的排法构成全集U, |U|=10!
  - 第1个数字不大于1的排法构成子集A(即所有以0或者1开头的排法), |A|=2\*9!
  - 最后一个数字不小于8的排法构成子集B(即所有以8或者9结束的排法), |B|=2\*9!
  - |A∩B|=2\*2\*8!
  - 题目要求的排法构成子集(~A∩~B)
  - $|(\sim A \cap \sim B)| = |U| |A \cup B| = |U| |A| |B| + |A \cap B| = 10! 4.9! + 4.8! = 2,338,560$

### 三个有限集合并集的计数



- 假设定义全集的三个子集A,B,C。则:
   |A∪B∪C|=|A|+|B|+|C|-|A∩B|-|A∩C|-|B∩C|+|A∩B∩C|
  - 证明:
  - $|A \cup B \cup C| = |A \cup B| + |C| |(A \cup B) \cap C|$
  - $= |A| + |B| |A \cap B| + |C| |(A \cap C) \cup (B \cap C)|$
  - =  $|A| + |B| |A \cap B| + |C| |(A \cap C)| |(B \cap C)| + |(A \cap B \cap C)|$
  - $= |A| + |B| + |C| |A \cap B| |A \cap C| |B \cap C| + |A \cap B \cap C|$

### 关于选课的例子



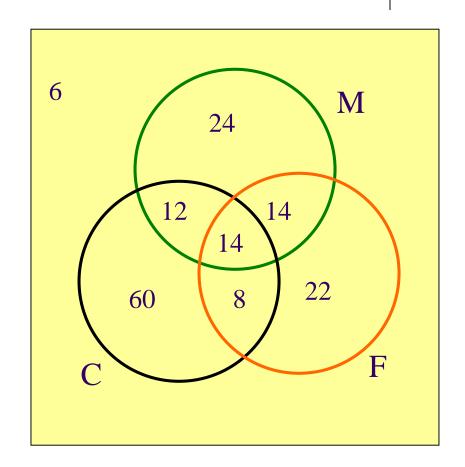
- 全班共有160个学生
  - 选数学课64人,选计算机课94人,选金融课58人
  - 选数学与金融的28人,选数学与计算机的26人,选 计算机与金融的22人
  - 三种课全选的14人。
- 问:这三种课都没选的是多少?只选一门计算机的有多少?

### 问题的解



- M-数学、C-计算机、F-金 融
- 包含-排斥原理
   |M∪C∪F|=|M|+|C|+|F| |M∩F|-|M∩C|-|C∩F|+
   |M∩C∩F|
   =64+94+58-28-26-22+14
   =154

因此, 6人未选课。 只选了计算机课的60人



### 容斥原理

#### (Principle of Inclusion and Exclusion)



假设 $A_1, A_2, ..., A_n$ 是n个有限集合,则它们的并集的元素个数是:

$$\bigcup_{i=1}^{n} A_{i} = S_{1} - S_{2} + S_{3} - \ldots + (-1)^{k-1} S_{k} + \ldots + (-1)^{n-1} S_{n}$$
 
$$\not \!\! \pm \psi \,, \quad S_{k} = \sum_{1 \leq i_{1} \leq i_{2} \leq \ldots \leq i_{k} \leq n} \!\!\! |A_{i_{1}} \cap A_{i_{2}} \cap \ldots \cap A_{i_{k}}| \qquad k = 1, 2, \ldots, n$$

### 例如: 4个子集的公式为:

 $|A_1| + |A_2| + |A_3| + |A_4|$ 

- $-(|A_1 \cap A_2| + |A_1 \cap A_3| + |A_1 \cap A_4| + |A_2 \cap A_3| + |A_2 \cap A_4| + |A_3 \cap A_4|)$
- $+ (|A_{1} \cap A_{2} \cap A_{3}| + |A_{1} \cap A_{2} \cap A_{4}| + |A_{1} \cap A_{3} \cap A_{4}| + |A_{2} \cap A_{3} \cap A_{4}|)$
- $|A_1 \cap A_2 \cap A_3 \cap A_4|$

### 容斥原理的证明



- 公式:  $\bigcup_{i=1}^{n} A_i = S_1 S_2 + S_3 ... + (-1)^{k-1} S_k + ... + (-1)^{n-1} S_n$
- 我们证明在上述公式中:
  - 并集中的元素在右边式子中恰好被计数1次
  - 设并集中对象a出现在m个集合中
  - 则它在在 $S_1$ 中被计数m 次,在 $S_2$ 中被计数  $C_2^m$ 次
    - 以n=4, m=3为例:
       |S₁|+ |S₂|+ |S₃|+ |S₄|
      - $-(|S_1 \cap S_2| + |S_1 \cap S_3| + |S_1 \cap S_4| + |S_2 \cap S_3| + |S_2 \cap S_4| + |S_3 \cap S_4|)$
      - $+ \; (|S_1 \cap S_2 \cap S_3| + |S_1 \cap S_2 \cap S_4| + |S_1 \cap S_3 \cap S_4| + |S_2 \cap S_3 \cap S_4|)$
      - $|S_1 \cap S_2 \cap S_3 \cap S_4|$

## 容斥原理的证明



- 公式:  $\bigcup_{i=1}^{n} A_i = S_1 S_2 + S_3 \dots + (-1)^{k-1} S_k + \dots + (-1)^{n-1} S_n$
- 我们证明并集中的元素在右边式子中恰好被计数1次
  - 设并集中对象a出现在m个集合中
  - 则它在在 $S_1$ 中被计数m 次,在 $S_k$ 中被计数  $C_k^m$ 次

### 容斥原理的证明



- 计数公式:  $\bigcup_{i=1}^{n} A_i = S_1 S_2 + S_3 \dots + (-1)^{k-1} S_k + \dots + (-1)^{n-1} S_n$
- 第二步: 满足1个或多个性质的元素恰好被计数0次:
  - 设对象a出现在m个集合中
  - a在S<sub>1</sub>中被计数  $C_1^m$ 次,S<sub>k</sub>中被计数恰好 $C_k^m$  次
  - 将上述分析带入计数公式可得:

$$C_1^m - C_2^m + ... + (-1)^{k-1} C_k^m + ... + (-1)^{m-1} C_m^m$$

• 该计算式值为1,因为当x=1时下式为0:

$$(1-x)^{m} = 1 - C_{1}^{m}x + C_{2}^{m}x^{2} + \dots + (-1)^{k}C_{k}^{m}x^{k} + \dots + (-1)^{m}C_{m}^{m}x^{m}$$

• a恰好被计数1次

### 埃拉托色尼的筛子

(Eratosthenes' Sieve)



• 用筛法求质数(以25以内的为例)

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

[2] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

[3] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

[5] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25



### 100以内有多少质数

- 100以内的任意合数必有不大于其平方根的质数为其因子。
   这样的质数只有4个: {2, 3, 5, 7}
- 设A<sub>2</sub>, A<sub>3</sub>, A<sub>5</sub>, A<sub>7</sub> 分别是可被相应质数整除的100以内大于1 的自然数的集合。则100以内质数的数量为:

$$N(\overline{A_2}\overline{A_3}\overline{A_5}\overline{A_7}) = 99 - \left\lfloor \frac{100}{2} \right\rfloor - \left\lfloor \frac{100}{3} \right\rfloor - \left\lfloor \frac{100}{5} \right\rfloor - \left\lfloor \frac{100}{7} \right\rfloor$$

$$+ \left\lfloor \frac{100}{2 \cdot 3} \right\rfloor + \left\lfloor \frac{100}{2 \cdot 5} \right\rfloor + \left\lfloor \frac{100}{2 \cdot 7} \right\rfloor + \left\lfloor \frac{100}{3 \cdot 5} \right\rfloor + \left\lfloor \frac{100}{5 \cdot 7} \right\rfloor + \left\lfloor \frac{100}{5 \cdot 7} \right\rfloor$$

$$- \left\lfloor \frac{100}{2 \cdot 3 \cdot 5} \right\rfloor - \left\lfloor \frac{100}{2 \cdot 3 \cdot 7} \right\rfloor - \left\lfloor \frac{100}{2 \cdot 5 \cdot 7} \right\rfloor - \left\lfloor \frac{100}{3 \cdot 5 \cdot 7} \right\rfloor + \left\lfloor \frac{100}{2 \cdot 3 \cdot 5 \cdot 7} \right\rfloor + \frac{4}{2 \cdot 3 \cdot 5 \cdot 7}$$

$$= 99 - 50 - 33 - 20 - 14 + 16 + 10 + 7 + 6 + 4 + 2 - 3 - 2 - 1 - 0 + 0 + 4$$

## 教材和作业

- 教材:
  - 2.4.5; 5.1.4; 7.5; 7.6
- 习题:
  - P120: 32; 35; 38 (第六版); P150: 3,17,19 (第七版)
  - P386: 8; 21 (第六版); P468: 8, 21 (第七版)

