组合数学第十三讲

授课时间: 2018年12月17日 授课教师: 孙晓明 记录人: 刘祥隆 史良

1 鸽笼原理(The Pigeonhole Principle)及其应用

鸽笼原理形象化的表述是，m只鸽子放入n个笼子中，若满足$m>n$，则一定存在一个笼子中至少含有两只鸽子。

定理 1 (鸽笼原理)**.** 设A是有限集，$|A|=m$。$A_i\subseteq A$ $(i=1,2,\cdots n)$，$m>n$，且满足

$$\bigcup_{i=1}^{n} A_i = A,$$

则必有正整数k $(1\le k\le n)$，使得$|A_k|\ge 2$。

定理 2 (鸽笼原理的一般形式)**.** 设A是有限集，$|A|=m$。$A_i\subseteq A$ $(i=1,2,\cdots n)$，且满足

$$\bigcup_{i=1}^{n} A_i = A,$$

则必有正整数k $(1\le k\le n)$，使得$|A_k|\ge\lceil\frac{m}{n}\rceil$。

证明 反证法。假设对任意的i，$|A_i|<\lceil\frac{m}{n}\rceil$，则$|A_i|\le\lceil\frac{m}{n}\rceil-1$。则$|A|=|\bigcup_{i=1}^{n}A_i|\le\sum_{i=1}^{n}|A_i|\le n\cdot(\lceil\frac{m}{n}\rceil-1)<m$，矛盾。最后一个不等号成立是因为，根据$\lceil\cdot\rceil$的定义，$\lceil\frac{m}{n}\rceil<\frac{m}{n}+1$。 □

例 1 从$\{1,2,3,\cdots,100\}$中任取51个整数，其中必有两个数互素。

证明 将$\{1,2,3,\cdots,100\}$按$\{1,2\},\{3,4\},\cdots,\{99,100\}$分为50组。依据鸽笼原理，总存在一组数，从中取到的数的个数不少于$\lceil\frac{51}{50}\rceil$，即为2；而由$\gcd(n,n+1)=1$知这两个数必互素，得证。

注意，本题只取50个数是不行的，因为可以取50个偶数$\{2,4,\cdots,100\}$。 □

例 2 从$\{1,2,3,\cdots,100\}$中任取51个整数，其中必有两个数a,b，满足$a\neq b$且$a\mid b$。

证明 参考分拆数部分$\mathrm{Odd}(n)=\mathrm{Diff}(n)$的证明。任意整数$n$可表示为$n=a\cdot 2^s$，其中$s$为非负整数，$a$为奇数。若$1\le n\le 100$，则$a\in\{1,3,5\cdots,99\}$。由于$a$只有50种取值，因此根据鸽笼原理，51个数中至少存在2个数，在它们的分解中a是相同的，此时两个数有整除关系。

注意，本题只取50个数是不行的，因为当取$\{51,52,\cdots,100\}$时，题述的性质不能满足。 □

例 3 将一个3×9的方格二染色，可以选出两行两列，使得它们交点上的方格同色。

证明 将3×1的方格二染色共有8种方法，，而3×9的方格共有9列3×1的方格，因此根据鸽笼原理，一定存在两列3×1的方格，二者的染色方案相同。现在只考虑这两列3×1的方格。根据鸽笼原理，对3×1的方格二染色，一定存在两行方格的颜色相同。因此我们选出了两行两列，它们交点上的方格同色，证毕。然而，下一个例题说明实际上并不需要这么多的方格。 □

例 4 将一个3×7的方格二染色，可以选出两行两列，使得它们交点上的方格同色。

证明 我们把对3×1的方格的8种染色方案记为$\{000,001,010,011,100,101,110,111\}$，染色方案$a_1a_2a_3$表示把$3\times 1$的方格的第$i$行染色为$a_i$，其中$a_i\in\{0,1\}, i=1,2,3$。现在把这8种方案分为6组：$\{000,001\},\{010\},\{011\},\{100\},\{101\},\{110,111\}$。根据鸽笼原理，在$3\times 7$的方格中，一定存在两列$3\times 1$的方格，选取了同一组的方案。若它们选取的是$\{010\},\{011\},\{100\},\{101\}$中的一组，则由例3的论证可得结论。若它们选择的是$\{000,001\}$，若二者的方案不同，则结论直接成立；若二者的方案相同，那么仍可使用例3的论证。最后，由对称性可知选择$\{110,111\}$时结论依然成立。

注意，在3×6的方格上，第1到6列分别染色为$001,010,100,110,101,011$，则不会出现题目要求的情况。 □

例 5 给定无向图$G=(V,E)$，一定存在两个顶点v_i,v_j，使得$\deg(v_i)=\deg(v_j)$。

证明 设图G有n个顶点，则$\deg(v)\in\{0,1,2,3,\cdots,n-1\}$。现将$\deg(v)$的取值范围分为$n-1$组：$\{0,n-1\},\{1\},\{2\},\{n-2\}$，根据鸽笼原理，存在两个顶点$v_i,v_j$，其取值范围落在同一组内。若它们不是落在第一组内，则命题成立。若它们同时落在第一组内，由于度为0的顶点和度为$n-1$的顶点不可能同时存在，因此它们的度必然同为0或者同为$n-1$，因此命题仍然成立。 □

例 6 若一个人用70天背110页书，且每天至少背一页，则必存在连续的若干天，他背书的总页数恰为29页。

证明 设第i天背书a_i页，以上问题等价于，若$a_1+a_2+\cdots+a_{70}=110$，$a_i\geq 1$且$a_i$均为整数,则存在$1\leq i\leq j\leq 70$，使得$a_i+a_{i+1}+\cdots+a_j=29$。

使用反证法，假设命题不成立。定义$S_i=a_1+a_2+\cdots+a_i$，表示前i天背书的总页数。由于$a_i\geq 1$，因此$1\leq S_1<S_2<\cdots<S_{70}=110$。对每个$S_i$加上29，得到$30\leq S_1+29<S_2+29<\cdots<S_{70}+29=139$。注意到，$\{S_i\}$之间两两不同，$\{S_i+29\}$之间也两两不同。根据假设，集合$A=\{S_i\}\cup\{S_i+29\}$中的任意两个数也不相同（若存在$1\leq i<j\leq 70$，使得$S_i+29=S_j$，那么$S_j-S_i=a_{i+1}+\cdots+a_j=29$，与假设矛盾）。那么有$|A|=140$。然而，$A$中整数的取值范围为$\{1,2,3,\cdots,139\}$，根据鸽笼原理，$A$中存在两个整数取值相同，矛盾。因此命题成立。 □

例 7 给定长度为n^2+1的实数列$\{a_1,a_2,\cdots,a_{n^2+1}\}$，数列的元素互不相同。证明该数列必然存在一个长度为$n+1$的单调递增子列或单调递减子列。即一定存在$n+1$个整数$1\leq i_1<i_2<\cdots<i_{n+1}\leq n^2+1$，使得$a_{i_1}<a_{i_2}<\cdots<a_{i_{n+1}}$或$a_{i_1}>a_{i_2}>\cdots>a_{i_{n+1}}$.

证明 假设该数列不存在长为$n+1$的单调递增子列。构造映射f，使得$f(a_i)$表示以a_i开始的最长递增子列的长度。由假设可知，对任意的$1\leq i\leq n^2+1$，均有$f(a_i)\in\{1,2,\cdots,n\}$。由于$f(a_i)$所有可能的取值仅有$n$个，据鸽笼原理，存在$\lceil\frac{n^2+1}{n}\rceil=n+1$个$a_i$对应的$f(a_i)$相同。设这$n+1$个$a_i$构成的子列为$a_{i_1},a_{i_2},\cdots,a_{i_{n+1}}$，则这一子列必定单调递减。论证如下：假设存在$k<l$使得$a_{i_k}<a_{i_l}$。由于存在以$a_{i_l}$打头的长度为$f(a_{i_l})$的递增子列，则$a_{i_k}$与这一子列一起构成了以$a_{i_k}$打头的长度为$f(a_{i_l})+1=f(a_{i_k})+1>f(a_{i_k})$的递增子列，矛盾。

注意，长度为n^2的数列不满足以上性质。考虑$n\times n$的矩阵$\{a_{ij}\}_{n\times n}$，其满足如下性质：若$i_1<i_2$，则$a_{i_1j}>a_{i_2j}$；若$j_1<j_2$，则$a_{ij_1}<a_{ij_2}$。将这一矩阵的元素按照从上到下、从左到右的顺序排成一个

数列，则这个数列中递增子列和递减子列的最大长度均不大于n。 □

例 8 在$\{1,2,\cdots,100\}$中取10个两两不同的数组成一个集合$S=\{a_1,a_2\cdots,a_{10}\}$，则存在$A,B\subseteq S$且$A\cap B=\emptyset$，满足 $\sum_{x\in A}x=\sum_{y\in B}y$

证明 若不包括空集，这样的S的子集个数为$2^{10}-1=1023$个。设T为S的子集，则T的子集和满足$55\le\sum_{x\in T}x\le 955$，可能的取值范围为$901<1023$。由鸽笼原理，一定存在两个子集$A$和$B$，使得它们的元素之和相同。若这两个子集相交，只需同时减去相交的部分，命题仍然成立。

容易知道，只取7个数时命题不成立，因为这时候可令$S=\{1,2,4,8,16,32,64\}$，它的任意两个子集的元素之和均不相同。通过计算机程序可以知道，当取8个数时，令$S=\{20,40,71,77,80,82,83,84\}$，它的任意两个子集的元素之和均不相同。可以证明，取9个数时，命题就成立了。

一般地，在$\{1,2,\cdots,n\}$中最多能选出多少个数组成一个集合S，使得S的任意两个子集的元素之和均不相同？事实上，已经证明存在常数c，使得$|S|\le\log n+c\log\log n$。 □

2 鸽笼原理与二次剩余

定理 3. *对于正整数n，存在整数x，y使得$n=x^2+y^2$，当且仅当任意的$4k+3$型素因子在n中出现偶数次。*

首先证明几个引理。

引理 4. *若对于正整数n_1,n_2，存在整数x_1,y_1,x_2,y_2，使得$n_1=x_1^2+y_1^2$，$n_2=x_2^2+y_2^2$，则存在x_3,y_3使得$n_3=n_1\cdot n_2=x_3^2+y_3^2$。*

证明 令复数$z_1=x_1-iy_1,z_2=x_2+iy_2$，则$z_3=z_1z_2=(x_1x_2+y_1y_2)+i(x_1y_2-x_2y_1)$。由$|z_1||z_2|=|z_1z_2|=|z_3|$得，如下恒等式成立，

$$(x_1^2+y_1^2)(x_2^2+y_2^2)=(x_1x_2+y_1y_2)^2+(x_1y_2-x_2y_1)^2。$$

令$x_3=x_1x_2+y_1y_2,y_3=x_1y_2-x_2y_1$，则有$n_3=n_1\cdot n_2=x_3^2+y_3^2$。 □

引理 5. *若p为$4k+1$型素数，则p可以被写成x^2+y^2的形式。*

证明 若p为$4k+1$型素数，则$(\frac{-1}{p})=1$，即存在z使得$z^2\equiv -1 \pmod p$。考虑所有形如$a+zb$的数，其中$a,b\in\{0,1,2\cdots\lfloor\sqrt{p}\rfloor\}$。显然这样的数共有$(\lfloor\sqrt{p}\rfloor+1)^2$个。由于$(\lfloor\sqrt{p}\rfloor+1)^2>p$，根据鸽笼原理，存在两个这样的数$a_1+zb_1,a_2+zb_2$在模$p$的意义下相等，即

$$a_1+zb_1\equiv a_2+zb_2 \pmod p。$$

移项得到，

$$(a_1-a_2)\equiv z(b_2-b_1) \pmod p。$$

两边同时平方得，

$$(a_1-a_2)^2\equiv(-1)(b_2-b_1)^2 \pmod p，$$

即，

$$(a_1-a_2)^2+(b_2-b_1)^2=0 \pmod p。$$

注意到，$a_1 - a_2 = 0$和$b_1 - b_2 = 0$不会同时成立，因此$(a_1 - a_2)^2 + (b_2 - b_1)^2 > 0$。另一方面，$(a_1 - a_2)^2 + (b_2 - b_1)^2 \leq 2(\lfloor\sqrt{p}\rfloor)^2 < 2p$。因此$(a_1 - a_2)^2 + (b_2 - b_1)^2 = p$，命题得证。 □

引理 6. *若q为$4k+3$型素数，且$q \mid n = x^2 + y^2$，则$q \mid x, q \mid y$。*

证明 利用反证法，假设$q \nmid y$。由已知条件，$x^2 \equiv -y^2 \pmod q$。根据假设，y在Z_q中存在逆元y^{-1}，因此有$(xy^{-1})^2 \equiv -1 \pmod q$，即$(\frac{-1}{q}) = 1$，与$q$是$4k+3$型素数矛盾。因此$q \mid y$，同理$q \mid x$。 □

以下证明原定理。

证明 设n的素因数分解为$n = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$。根据引理4，可以假设$k_i \in \{0, 1\}$。由2可以写成$2 = 1 + 1$，以及引理5，可以删去$n$中的素因子2和所有$4k+1$型的素因子，故可以假设$n = q_1 q_2 \cdots q_l$，其中$q_i$为$4k+3$型的素因子。

“当”：当任意的$4k+3$型素因子在n中出现偶数次时，$l = 0$，即$n = 1$，显然n可以进行写成平方和的形式。

“仅当”：若存在$4k+3$型的素因子在n中出现奇数次，则$l > 0$。利用反证法，假设存在正整数x, y，使得$n = x^2 + y^2$，由于$q_i \mid n$，根据引理6，$q_i \mid x, q_i \mid y$。由于q_i为素数，有$n = q_1 q_2 \cdots q_l \mid x, n = q_1 q_2 \cdots q_l \mid y$，矛盾。 □