Pigeonhole Principle and its applications

Mark Kirichev

July 13, 2022

Abstract

The pigeonhole principle is a powerful tool that is found in all kinds of combinatorial problems. Using the principle we can easily construct contradiction proofs to almost any problem that requires some kind of counting. Moreover, the Pigeonhole principle is so useful that it has applications in combinatorial geometry, number theory, algebra, and even statistics. For example, thanks to it we know that in New York City, there's at least 2 people with the same number of hairs! We'll go over Erdos-Szerekes's Upper bound for Ramsey numbers, Kronecker Dirichet's Approximation theorems, and some interesting Olympiad problems from the Moscow Math Olympiad, Ireland Math Olympiad, USAMO, and IMO.

1 Introduction to Pigeonhole Principle

General Introduction and a few very basic examples. Maybe a visual representation fron the internet.

2 Problems

2.1 Problem 1

Problem 1G - Lint and Wilson
Show that a finite simple graph with more than one vertex has at least two vertices with the same degree.

2.2 Problem 2

Problem 1: Erdos-Szerekes Upper bound - Po-Shen Loh, June 2012
The Ramsey Number $R(s,t)$ is the minimum integer n for which every red-blue coloring of the edges of K_n contains a completely red Ks or a completely blue K_t. Prove that:

$$R(s,t) \leq \binom{s+t-2}{s-1}$$

2.3 Problem 3

Problem 2: Moscow Math Olympiad - Po-Shen Loh, June 2012

Show that any convex polyhedron has two faces with the same number of edges.

2.4 Problem 4

Problem 4: Ireland Math Olympiad 2012 - Po-Shen Loh, June 2012

The numbers $1, 2, ..., 4n^2$ are written in the unit squares of a $(2n)$ x $(2n)$ array, $3 \leq n$. Prove that there exist $n+1$ columns in the array such that in each of them any number is less than the sum of the remaining $2n$ - 1 numbers in that column.

2.5 Problem 5

USAMO 1976

Every square in a 4 x 7 array is colored either white or black. Show that there always is a monochromatic "constellation" consisting of the 4 corners of an axis-parallel rectangle.

2.6 Problem 6

Problem 6: Erdos-Szekeres - Po-Shen Loh, June 2010

Prove that every sequence of n^2 distinct numbers contains a subsequence of length n which is monotone (i.e. either always increasing or always decreasing).

2.7 Problem 7

Kronecker's theorem

Let x be an irrational number, and let $x_n = nx$ be the fractional part of nx. Show that the sequence $x_1, x_2, ...$ is dense in the interval $[0, 1)$. This means that for every real number $r \in [0, 1)$, and every $\epsilon > 0$, there is some n such that xn is within ϵ of r.

2.8 Problem 8

Example 9 [IMO Shortlist 2001, C6] - Olympiad Combinatorics, Sriram

For a positive integer n define a sequence of zeros and ones to be balanced if it contains n zeros and n ones. Two balanced sequences a and b are neighbors if you can move one of the $2n$ symbols of a to another position to form b. For instance, when n = 4, the balanced sequences 01101001 and 00110101 are neighbors because the third (or fourth) zero in the first sequence can be moved to the first or second position to form the second sequence. Prove that there is a set **S** of at most balanced sequences such that every balanced sequence is equal to or is a neighbor of at least one sequence in **S**.

2.9 Problem 9

Example 11 [Russia 2011, adapted] - Olympiad Combinatorics, Sriram

There are $N > n^2$ stones on a table. **A** and **B** play a game. **A** begins, and then they alternate. In each turn a player can remove k stones, where k is a positive integer that is either less than n or a multiple of n. The player who takes the last stone wins. Prove that **A** has a winning strategy.

2.10 Problem 10

Problem 13 - Po-Shen Loh, June 2010

The Fibonacci numbers are defined by $F_1 = F_2 = 1$ and $F_n = F_{n1} + F_{n2}$ for $3 \leq n$. If p is a prime number, prove that at least one of the first $p + 1$ Fibonacci numbers must be divisible by p.

3 Solutions

3.1 Problem 1

Assume the graph **G** has n vertices. The possibilities for the degree are $0, ..., n - 1$. However, **G** cannot have one vertex of degree 0 and one vertex of degree $n - 1$, because these two vertices would need to be adjacent to satisfy degree $n - 1$. Using the pigeonhole principle, we must pick the n degrees (one for each vertex), from a set of $n - 1$ answers (since we cannot have both 0 and $n - 1$). Hence, one degree must be repeated.

3.2 Problem 2

Observe that $R(s, t) \leq$R(s 1, t) + R(s, t 1), because if we have that many vertices, then if we select one arbitrary vertex WLOG, then it cannot simultaneously have < R(s - 1, t) red neighbors and < R(s, t - 1) blue neighbors, so we can inductively build either a red K_s or a blue K_t. But we have:

$$\binom{(s-1)+t-2}{s-2} + \binom{s+(t-1)-2}{s-1} = \binom{s+t-2}{s-1}$$

because in Pascal's Triangle the sum of two adjacent guys in a row equals the guy directly below them in the next row.

3.3 Problem 3

Consider the dual graph, where faces are vertices and adjacent faces give edges. Every graph has two vertices of equal degree, which we found in Problem 1.

3.4 Problem 4

Suppose for contradiction that the first n columns all have that the largest number is at least the sum of all other numbers in that column. Let **B** be the sum by taking the largest number in each of those columns, and let **A** be the sum by taking all but the largest number in each of those columns. Then we have $\mathbf{A} \leq \mathbf{B}$. However, the smallest **A** can be is if it is the first $n(2n1)$ numbers, and the largest **B** can be is if it is the last n numbers. Hence, we have the following:

$1 + 2 + 3 + ... + n(2n-1) \leq (4n^2 - n + 1) + (4n^2 - n + 2) + ... + 4n^2$
$1 + 2 + 3 + ... + n(2n-1) \leq (4n^2 - n) + 1 + (4n^2 - n) + 2 + ...(4n^2 - n) + n$
$1 + 2 + 3 + ... + n(2n-1) \leq n(4n^2 - n) + (1 + 2 + 3 + ... + n)$

By using Gaussian summation on both sides, we can get that:

- $$\sum_{i=1}^{n(2n-1)} i = \frac{n(2n-1)[n(2n-1)+1]}{2}$$
- $$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

When we substitute those results into our inequality, we get:
$\frac{n(2n-1)[n(2n-1)+1]}{2} \leq n(4n^2 - n) + \frac{n(n+1)}{2}$
$n(2n-1)[n(2n-1)+1] \leq 8n^3 - 2n^2 + n(n+1)$
$n(2n-1)[n(2n-1)+1] \leq n(8n^2 - 2n + n + 1)$
$(2n-1)[n(2n-1)+1] \leq 8n^2 - n + 1$
$n(2n-1)^2 + (2n-1) \leq 8n^2 - n + 1$
$n(4n^2 - 4n + 1) + 2n - 1 \leq 8n^2 - n + 1$
$4n^3 - 4n^2 + n + 2n - 1 \leq 8n^2 - n + 1$
$4n^3 - 4n^2 + 3n - 1 \leq 8n^2 - n + 1$
$4n^3 \leq 12n^2 - 4n + 2$ $2n^3 \leq 6n^2 - 2n + 1$

It's obvious that the functions on bots sides of the inequality are continuous, growing as $n \to \infty$, and that the LHS will dominate the RHS. This happens when $3 \leq n$ (this can be checked pretty easily in a number of ways). Thus, we're done!

3.5 Problem 5

Key observation: $7 = \binom{4}{2} + 1$. First, suppose that some column contains 3 or more of the same color, say black. WLOG, they are in the first 3 rows. Then, for the other 6 columns, there cannot ever be 2 black in the first 3 rows, i.e., there must be at least 2 white in the first 3 rows. The number of ways to do this is:

$3 while \to 1$
$2 white \to 3$

Therefore, by the fifth column, there is repetition, and this gives the 4 corners. Thus, in every column there are exactly 2 of each color. But then after $\binom{4}{2}$ columns, there is a repeat, so we get the 4 corners again!

3.6 Problem 6

For each of the n^2 indices in the sequence, associate the ordered pair (x, y) where x is the length of the longest increasing subsequence ending at x, and y is the length of the longest decreasing one. All ordered pairs must obviously be distinct. But if they only take values with $x, y \in$1, . . . , n 1, then there are not enough for the total n^2 ordered pairs. Thus n appears somewhere, and we are done!

3.7 Problem 8

Call such a set **S** a dominating set. Our idea is to partition the set of $\binom{2n}{n}$ balanced sequences into (n+1) classes, so that the set of sequences in any class form a dominating set. Then we will be done by the pigeonhole principle, since some class will have at most $\frac{1}{n+1}\binom{2n}{n}$ balanced sequences. To construct such a partition, for any balanced sequence A let f(A) denote the sum of the positions of the ones in $A(\mod (n+1))$. For example, f(100101) $\equiv 1+4+6(\mod 4) \equiv 3 \mod 4$. A sequence is in class i if and only if f(A) $\equiv$i $\mod (n+1)$. It just remains to show that every class is indeed a dominating set, that is, for any class C_i and any balanced sequence A not in C_i, A has a neighbor in C_i. This isn't difficult: if A begins with a one, observe that moving this one immediately $\equiv f$(A) $+ kmod(n+1)$. Hence simply choose $k \equiv I - f(A) \mod (n+1)$, and then by shifting the first one to the right of the k-th zero we end up with a sequence B satisfying f(B) $\equiv$i$mod(n+1)$. Hence B is a sequence in C_i. The case when A begins with a zero is similar. Thus each class is indeed a dominating set and we are done by the first paragraph.

References

[1] Po-Shen Loh, June 2012, https://www.math.cmu.edu/ ploh/olympiad.shtml.

[2] Po-Shen Loh, June 2010, https://www.math.cmu.edu/ ploh/olympiad.shtml.

[3] Olympiad Combinatorics, Pranav A. Sriram, Aug 2014.

[4] A Course in Combinatorics, J. H. van Lint R. M. Wilson, 2nd edition, 2001.