Pigeonhole Principle

Rajiv Raman

October 21, 2023

Abstract

These notes provide a quick background on the pigeonhole principle.

1 Introduction

The pigeonhole principle (PHP) is the following simple observation that if there are $n+1$ pigeons and only n pigeonholes, then some pigeonhole contains more than one pigeon. This simple observation finds utility in many branches of mathematics, and is important enough to warrant a name. In these notes we will see some interesting applications of the pigeonhole principle.

The principle is simple enough that to appreciate its utility we should look at examples.

Example 1

If more than 12 people are assembled in a room, then there must be two people whose birthday is on the same month.

Here, there are 12 pigeonholes corresponding to the 12 months and the people correspond to the pigeons. The result follows directly from the pigeonhole principle.

Example 2

If there are n married couples, how many among the $2n$ people must we choose before we are sure that there is at least one married couple we have chosen?

Here the pigeonholes correspond to the n married couples. If we choose $n+1$ people, then by the pigeonhole principle, two of them must fall into the same pigeonhole, and in other words, they are married.

Let us now look at a slightly more challenging example.

Example 3

Given any set of n integers $a_1, \ldots, a_n$, prove that there must exist $0 \leq k < \ell \leq n$ s.t. $\sum_{j=k+1}^{\ell} a_j$ is divisible by n.

Here, it is not immediately clear what corresponds to the pigeons and what corresponds to the pigeonholes. In such cases, some ingenuity is required in constructing the sets corresponding to the pigeons and the pigeonholes.

For $t = 1, \ldots, n$ let $S_t = \sum_{j=1}^{t} a_i$. If any one of these n partial sums are divisible by n, we are done. So we can assume that $S_i \equiv s \mod n$, where $s \in \{1, \ldots, n-1\}$. Here we have a potential application of the pigeonhole principle - There are n partial sums $S_1, \ldots, S_n$ but only $n-1$ possible reminders. Therefore by PHP, there exist k and ℓ s.t. $S_k \mod n \equiv S_\ell \mod n$. But, this implies $S_k - S_\ell \equiv 0 \mod n$, and we have found the desired k and ℓ.

Example 4

A chess master wants to prepare for a tournament. There are 11 weeks left. In order to stay in the zone, (s)he wants to play at least one game a day, but in order not to get too exhausted doesn't play more than 12 games in any 7 day period. Prove that there is a sequence of consecutive days where (s)he plays exactly 21 games.

There are 77 days. Let $a_1, \ldots, a_{77}$ be the number of chess games played on each of the days. As in the previous problem, it is not apparent what corresponds to the pigeons and what corresponds to the pigeonholes. As before, let $S_k = \sum_{j=1}^{k} a_j$. The sequence is strictly increasing since the chess player plays at least one game a day. Since (s)he plays at most 12 games in any 7 day period, $S_{77} \leq 12 \times 11 = 132$. Thus, we have

$$1 \leq S_1 < \ldots < S_{77} \leq 132$$

We construct a new sequence $S'_1, \ldots, S'_{77}$, where $S'_i = S_i + 21$. From the previous argument it follows that

$$22 \leq S_1 < \ldots < S_{77} \leq 153$$

Hence, the numbers $\{S_i : i = 1, \ldots, 77\} \cup \{S'_i : i = 1, \ldots, 77\}$ lie in the range $\{1, \ldots, 153$. There are $77 + 77 = 154$ numbers in the sequence. Now we see a potential way to apply the pigeonhole principle - there are 154 elements that lie in the range $1, \ldots, 153$. Therefore, by PHP, two of them must be equal. No two S'_is can be equal as the sequence is strictly increasing. By the same argument, no two S'_i can be equal. Therefore, there exist $1 \leq j < k \leq 77$ s.t $S_j = S'_k$. This implies that $S_k - S_j = 21$. Thus, (s)he plays exactly 21 games between days k and ℓ.

Exercise 1

From the integers $1, \ldots, 200$ we choose 101 integers. Show that among the chosen integers there are two such that one is divisible by the other.

Hint: We can write each integer $n_i \in \{1, \ldots, 200\}$ in the form $2^i a_i$, where a_i is odd.

2 Strong Pigeonhole principle

Theorem 1

Let $q_1, \ldots, q_n$ be positive integers. If $q_1 + \ldots + q_n - n + 1$ balls are distributed into n boxes $B_1, \ldots, B_n$, then either B_1 receives q_1 balls, or B_2 receives q_2 balls,…, or B_n receives q_n balls.

The standard PHP follows by setting $q_i = 2$ for each $i = 1, \ldots, n$.

Proof. By contradiction. If each B_i contains at most $q_i - 1$ balls, then we have in total $(q_1 - 1) + (q_2 - 1) + \ldots + (q_n - 1) = q_1 + \ldots + q_n - n$ balls. But, we have one more ball. □

Corollary 1. *If $n(r-1)+1$ objects are distributed into n boxes, then at least one box contains at least r objects.*

We can view this alternatively as an *averaging argument*. If the average of n integers $m_1, \ldots, m_n$ is greater than $r-1$, i.e. $(m_1 + \ldots + m_n)/n > r-1$, then for at least one i, $m_i \geq r$. Indeed, if all the m_i's are at most $r-1$, then their average is at most $n(r-1)/n$, but the average is greater than $r-1$. Since the m_i's are integers, one of them must be at least r.

Similarly, if the sum of n integers $m_1, \ldots, m_n$ is less than $r+1$, then one of them is at most r.

Example 5

Two disks, one smaller than the other are each divided into two hundred congruent sectors. In the larger disk, 100 of the sectors are painted blue and the remaining are painted red. In the smaller disk, each sector is painted red or blue with no stipulation on the number of sectors colored blue or red.

The smaller disk is placed on the larger disk so that their centers align. Show that there is a rotation of the smaller disk so that the number of sectors of the smaller disk whose color aligns with that of the larger disk is at least 100.

Number the sectors of each disk $1, \ldots, 200$. We try all possible alignments of the smaller disk. There are 200 possible alignments. Since the larger disk has 100 red and 100 blue sectors, for a fixed sector of the smaller disk, there are 100 rotations where the color of this sector matches that of the larger disk. Therefore, over all alignments and all sectors of the smaller disk, there are $100 \times 200 = 20,000$ matches. Since there are 200 sectors of the smaller disk, on average a sector of the smaller disk matches 100 sectors of the larger disk. Therefore, there is a sector of the smaller disk whose alignment with the larger disk is at least 100.

Another way to view this is as follows. Construct a 200×200 matrix whose entries are 0 or 1. The columns correspond to the sectors of the smaller disk. The rows correspond to the 200 circular alignments of the smaller disk. The conditions of the problem guarantees that in each column there are at least 100 1's. Therefore, overall there are at least $100 \times 200 = 20,000$ 1's. Therefore, there must be a row with at least 100 1's.

2.1 Erdös-Szekeres Theorem

Now we present a famous result of Erdös and Szekeres.

Given a sequence $a_1, \ldots, a_k$ of elements, a subsequence is a sequence of elements $a_{i_1}, a_{i_2}, \ldots, a_{i_\ell}$, where $1 \leq i_1 < i_2 < \ldots < i_\ell \leq k$. A sequence if non-decreasing if $a_{i_r} \leq a_{i_s}$ whenever $r < s$ and is non-increasing if $a_{i_r} \geq a_{i_s}$ for all $r < s$.

Theorem 2

In any sequence of $n^2 + 1$ integers, there either a non-decreasing subsequence of length $n + 1$ or a non-increasing subsequence of length $n + 1$.

Proof. We prove that if there is no non-decreasing subsequence of length $n + 1$, there must be a non-increasing subsequence of length $n + 1$. Let the sequence be $a_1, \ldots, a_{n^2+1}$. Let s_k denote the length of the longest non-decreasing subsequence that ends at a_k. By assumption, $s_k \leq n$ for each $k = 1, \ldots, n^2 + 1$. Therefore, we have $n^2 + 1$ numbers $s_1, \ldots, s_{n^2+1}$ that lie in the range $\{1, \ldots, n\}$. Therefore, by the strong form of the pigeonhole principle, there must be $n+1$ that have the same value. That is, there is a subsequence $a_{i_1}, \ldots, a_{i_{n+1}}$ such that the longest non-decreasing subsequence ending at a_{i_j} have the same length. We must have $a_{i_j} \geq a_{i_{j+1}}$ for each $j = 1, \ldots, n+1$. Otherwise, we can combine the non-decreasing subsequence ending at s_{i_j} with $s_{i_{j+1}}$ to obtain a longer non-decreasing subsequence ending at $s_{i_{j+1}}$.

Therefore, $a_{i_1} \leq \ldots \leq a_{i_{n+1}}$ and is a non-increasing subsequence of length $n + 1$. $\square$

Exercise 2

Give examples of subsequence of length n^2+1 that contain no non-decreasing subsequence of length $n + 1$ but only a non-increasing subsequence of length $n + 1$.