

PIGEONHOLE PRINCIPLE

Pigeonhole principle is a fundamental but powerful tool in combinatorics. Unlike many other strong theorems, the principle itself is exceptionally simple. Unless you have looked into it thoroughly, it is hard to have a glimpse of its elegance and useful applications in mathematics.

1. The simple but elegant statement

Pigeonhole Principle states a trivial fact as follows:

Theorem 1.1. (Pigeonhole Principle)

When n pigeons are put into k pigeonholes, there exists at least 1 pigeonhole containing not less than $\lceil \frac{n}{k} \rceil$ pigeons and at least 1 pigeonhole containing not more than $\lfloor \frac{n}{k} \rfloor$ pigeons.

➤ $\lceil x \rceil$ and $\lfloor x \rfloor$ represent the smallest integer not less than x and the greatest integer not greater than x respectively. They are called ‘ceiling function’ and ‘floor function’. For example, $\lceil \pi \rceil = 4$, $\lceil 2 \rceil = 2$, $\lfloor \pi \rfloor = 3$, $\lfloor 2 \rfloor = 2$.

Proof. We prove the first half of the statement. Assuming the contrary, then all pigeonholes contain not more than $\lceil \frac{n}{k} \rceil - 1$ pigeons. Recall that $\lceil x \rceil - 1 < x$ for all real numbers x . Therefore, there are less than $k \times (\lceil \frac{n}{k} \rceil - 1) < k \times \frac{n}{k} = n$ pigeons. Contradiction. The remaining part can be proved similarly by using the inequality $\lfloor x \rfloor + 1 > x$.

Q.E.D.

Note that the theorem only mentions the existence of a pigeonhole satisfying some desired properties. Therefore, it can be used in non-constructive proofs or solutions usually. (Of course, there are exceptions. It will be illustrated in examples later.)

Besides, as seen in the statement, there are 2 unusual terms: ‘pigeons’ and ‘pigeonholes’. They are not readily available in usual mathematical problems. Therefore, we have to create them suitably in order to adopt the theorem. Only suitable constructions can lead to a solution. Here are some simple examples illustrating the theorem directly:

Example 1.1.

51 numbers are chosen from the integers between 1 and 100 inclusively. Prove that 2 of the chosen integers are consecutive.

Solution.

Since the existence of consecutive integers is the main theme of the problem, it is natural to form pigeonholes using consecutive integers. We partition the 100 integers into 50 pairs of consecutive integers as pigeonholes: $\{1, 2\}, \{3, 4\}, \{5, 6\}, \dots, \{99, 100\}$. Let the 51 chosen integers be pigeons. By Pigeonhole Principle, when we choose 51 integers, there is at least 1 pigeonhole (a pair of integers) containing $\lceil \frac{51}{50} \rceil = 2$ pigeons (chosen integers). Therefore, there are 2 consecutive integers among the 51 chosen integers.

Example 1.2.

10 integers are chosen from 1 to 100 inclusively. Prove that we can find 2 disjoint non-empty subsets of the chosen integers such that the 2 subsets give the same sum of elements.

Solution.

Note that we are asked to find 2 subsets giving the same sum of elements. Therefore, we may let 'the possible sum of elements' be pigeonholes. Although we do not know what the 10 chosen numbers are, we know the range of the sums. Clearly, the sum must be an integer. Moreover, the sum must be at least 1 (since all chosen integers are positive) and at most $91+92+\dots+100=955$. Therefore, we may set 955 pigeonholes as 1, 2, 3, ..., 955 and the subsets of the chosen numbers as pigeons. Obviously, there are $2^{10}-1=1023$ pigeons (non-empty subsets). Therefore, there is at least 1 pigeonhole with $\lceil \frac{1023}{955} \rceil = 2$ pigeons. i.e. 2 subsets giving the same sum of elements. Note that any one of them cannot be a subset of another. Otherwise, they must not have the same sum of elements. If the 2 subsets are disjoint, we are done. If they have common elements, we may remove the common elements from the 2 subsets. It reduces the sums by the same amount (namely the sum of common elements), so the 2 new subsets give the same sum again.

In Example 1.1 and 1.2, since we are asked to prove the existence of several objects satisfying some desired properties (say, 2 subsets giving the same sum), it is natural to set pigeonholes as the different statuses of the properties (say, the different sums) and the pigeons as the objects (say, the subsets). Examples 1.1 and 1.2 illustrate these direct applications. Besides, problems about colouring may also be solved using Pigeonhole Principle.

Example 1.3. (Ramsey Number $R(3, 3)$)

6 points are drawn on a plane. All edges between every 2 points are painted in red or blue. Prove that there is a monochromatic triangle (i.e. a triangle with its 3 edges painted in the same colour).

Solution.

The problem seems irrelevant to Pigeonhole Principle. In fact, they have some connections. What are the pigeonholes and pigeons? They are not explicitly stated in the problem. However, we may think that the triangles are formed by edges. If one of the colours, say red, dominate the edges of the graph, it should be easy to find a red triangle. In other words, if the number of edges of one of the colours is guaranteed, the problem may be solved. Note that there are $C_2^6 = 15$ edges in the graph, painted in 2 colours, so we may think simply that there are at least $\lceil \frac{15}{2} \rceil = 8$ edges painted in the same colour, say red, by Pigeonhole Principle. However, it does not guarantee a red triangle (See Figure 1.1). Therefore, this setting of pigeonholes and pigeons fails.

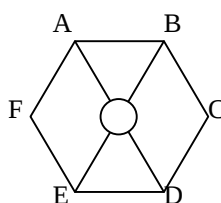


Figure 1.1: No triangle is formed with 8 edges.

(Note that the point in the centre should be ignored).

Nevertheless, we may find another way to define pigeonholes from Figure 1.1. If the unpainted edges are painted in blue, we can see there are many blue triangles despite the smaller number of blue edges. The main reason is that many blue edges meet at the same point. Inspired from this situation, we may focus on one point only and look into the edges incident to it. Pick any point among the 6 points, say A. There are 5 edges incident to A. By Pigeonhole Principle, there are $\lceil \frac{5}{2} \rceil = 3$ edges painted in the same colour, say red. Without loss of generality, let AB, AC and AD be red. Next, if any one of BC, BD and CD is red, then there exists a red triangle. Otherwise, BC, BD and CD form a blue triangle. Therefore, in both cases, a monochrome triangle appears.

From this example, we see that the way to construct pigeonholes is essential. Since the pigeonholes may not be explicit, there may be many ways to construct pigeonholes. However, only nice ways of construction of pigeonholes lead to a solution. Therefore, we are going to look into the constructions of pigeonholes in the following paragraphs.

➤ If you know Ramsey's number, you will know that this proof guarantees $R(3, 3) \leq 6$.

Exercise.

1. 11 integers are randomly chosen. Prove that two of them have the same unit digit.
2. There are some people (more than 1 person) in a party. Prove that 2 of them have the same number of friends in the party. (It is assumed that friendship is mutual. That is if A is a friend of B , then B is a friend of A .)
3. 9 people are in a club. Each of them can play one of the games among Bridge, Hearts and Mahjong. Prove that they can play at least one of the mentioned games. (Note that all 3 games require 4 players.)
4. There are 17 mathematicians and 3 official languages. Every pair of mathematicians communicate in one of the official languages. Prove that there are 3 mathematicians communicating in the same language pairwise.

2. How to construct pigeonholes?

Pigeons and pigeonholes are sometimes abstract. Therefore, drawing out some cases is the best way to find clues for the construction of the pigeonholes. From the trials, you will see how the pigeonholes should be constructed.

Example 2.1.

There are 5 points in a square of side length 2. Prove that there exist 2 of them having a distance not more than $\sqrt{2}$.

Solution.

After reading the question, we should try to think of the reason for the desired result.

Since the unit square is bounded, in order to maximize the minimal distance between the points, we should put the points as 'far' as we can. The most intuitive way to do so is putting them at corners. However, there are only 4 corners, so one point should be placed so that it is 'quite' far from the 4 corners. Clearly, the centre of the square is such a point. It is quite 'obvious' that the minimal distance between the points is the greatest now. (See Figure 2.1)

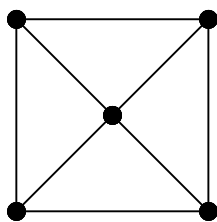


Figure 2.1: The 5 points are at corners and the centre.

However, how can we claim the minimum distance cannot be greater than $\sqrt{2}$ in all circumstances? Actually, the figure tells us the answer. If we divide the square into 4 unit squares, then any two points in the same unit square are at most $\sqrt{2}$ apart. (See Figure 2.2) By Pigeonhole Principle, when 5 points are put into the squares, two of them must be in the same unit square. These two points are what we want to find.

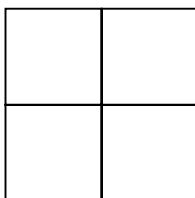


Figure 2.2: The square is divided into 4 unit squares. There are always 2 points in 1 unit square.

- The aid of the figure of the ‘just satisfaction of the requirement’ is important. It inspires us to think how we should divide the square. For example, you will know that dividing the square into 4 big triangles (using the 4 sides as the bases) is useless from Figure 2.1.
- If you check the Principle carefully, you will find that it is assumed implicitly that each pigeon are put in one pigeonhole only. Therefore, strictly speaking, some of the 4 unit squares should have lost a part of their common boundaries with other unit squares so as to be consistent with this implicit requirement of the theorem.

Example 2.2.

27 points are aligned so that each row has 9 points and each column has 3 points. (A column is perpendicular to a row.) Each point is painted in red or blue. Prove that there exists a monochromatic rectangle (i.e. 4 vertices are of the same colour) with its sides parallel to the rows and columns.

Solution.

Try to draw the points to see how they ‘behave’. If you try to draw and paint them randomly, you will probably find that some columns are exactly the same. Actually, this is the way to the solution.

Every point may have 2 colours. Therefore, every column containing 3 points has $2^3 = 8$ colouring schemes. By Pigeonhole Principle, there are at least $\lceil \frac{9}{8} \rceil = 2$ columns painted exactly the same. In each of these 2 columns, by Pigeonhole Principle again, we have at least 2 points in the column painted in the same colour. Obviously, these 4 points having the same colour in these 2 columns form a rectangle with the desired properties.

- Drawing the points to see how they behave is important. It helps us find the way to construct pigeonholes.
- Actually, 21 points aligned in 7 rows and 3 columns are enough to find such a rectangle.

Example 2.3.

76 points are aligned so that each row has 19 points and each column has 4 points. (A column is perpendicular to a row.) Each point is painted in red, blue or yellow. Prove that there exists a monochromatic rectangle (i.e. 4 vertices are of the same colour) with its sides parallel to the rows and columns.

Solution.

Since there are $3^4 = 81$ colouring schemes for a column, it is impossible to adopt the arguments in Example 2.2 here. However, we see that it is not necessary to have 2 identical columns for the existence of a monochromatic rectangle. We only need two columns, each has 2 points in the same colour and in the same rows. Therefore, we can reduce our pigeonholes as follows:

Denote R, B, Y as red, blue and yellow respectively. Let the 19 columns be pigeons and $(X, \{a, b\})$ be pigeonholes where $X = R, B, Y$ and $1 \leq a < b \leq 4$. Note that there are 4 points in each column. By Pigeonhole Principle, at least $\lceil \frac{4}{3} \rceil = 2$ of them are in the same colour. For each column, if the points at column a and b are of color X , we put it into the pigeonhole $(X, \{a, b\})$. (If it has more than 1 pair of points in the same colour, just choose any one of the pigeonholes.) There are $3 \times C_2^4 = 18$ pigeonholes. By Pigeonhole Principle, $\lceil \frac{19}{18} \rceil = 2$ of them are in the same pigeonhole, say $(X_1, \{a_1, b_1\})$, then we have a rectangle in colour X_1 in columns a_1, b_1 .

Example 2.4.

There is a sequence of 100 integers. Prove that there is a sequence of consecutive terms such that the sum of these terms is divisible by 99.

Solution.

Note that a number divisible by 99 can be expressed as a difference of two numbers which have the same remainder when divided by 99. Therefore, this problem is equivalent to finding 2 sequences of consecutive terms where one of them is a subsequence of the other and they give the same sum of elements.

Having this concept, it is natural to think of the following construction: Let S_k be the subsequence formed by the first k terms of the original sequence and A_k be its sum of elements. Note that for any $1 \leq i < j \leq 100$, S_i is contained in S_j . Since the remainder when A_k is divided by 99 has at most 99 possibilities. By Pigeonhole Principle, among $A_1, A_2, \dots, A_{100}$, there are at least $\left\lceil \frac{100}{99} \right\rceil = 2$ of them having the same remainder when divided by 99, say A_i and A_j . Then the sequence obtained by removing the first i th term from S_j (i.e., the subsequence from the $(i+1)$ th term to the j th term) has a sum of elements divisible by 99.

Example 2.5.

51 integers are chosen from $\{1, 2, 3, \dots, 100\}$. Prove that there are 2 distinct chosen integers such that one is a multiple of the other.

Solution.

If we would like to use Pigeonhole Principle to solve this problem, it seems that we have to partition $\{1, 2, 3, \dots, 100\}$ into 50 subsets such that every 2 elements of a subset have a relation that one is a multiple of the other.

In order to do this, we may consider the singletons of odd numbers $\{1\}, \{3\}, \dots, \{99\}$. Next, we append these singletons. For each odd number k , we put $2k, 4k, 8k, \dots, 2^i k, \dots$ into the set $\{k\}$. Each such subset is a pigeonhole. (i.e. they are $\{1, 2, 4, \dots, 64\}, \{3, 6, 12, \dots, 96\}, \{5, 10, \dots, 80\}, \dots, \{97\}, \{99\}$). By Pigeonhole Principle, $\left\lceil \frac{51}{50} \right\rceil = 2$ of them are in the same subset, so the larger one is a multiple of the smaller one.

Exercises.

1. 9 points are in a unit square. Prove that there is a triangle formed by 3 of the points has an area not more than $\frac{1}{8}$.

2. There are 15 subjects in a school. Every student takes 4 subjects out of them. Given that the school has 18 students, prove that 2 of the students have 2 common subjects.
3. 1502 people are seated in a row of 2002 chairs. Prove that there are 3 consecutive non-empty chairs.
4. A child drinks at least 1 bottle of milk a day. Given that he has drunk 700 bottles of milk in a year of 365 days, prove that for he has drunk exactly 29 bottles in some consecutive days.
5. 14 integers are chosen from $\{1, 2, \dots, 28\}$. Prove that there exist four of the chosen integers which can be split into two groups of two with the same sum.
6. Consider the sequence $\{\sqrt{3}, 2\sqrt{3}, 3\sqrt{3}, \dots\}$. Prove that some of the terms have a mantissa of less than 0.01 when written in decimal representation. (Consider the arguments used in Example 2.4.) Besides, can the number 0.01 be replaced by a smaller positive number? [Note: the mantissa of a real number means its 'decimal part'. For example, the mantissa of π is 0.1415926....]

3. Finding the bound

In the sessions above, all problems only require a proof for the statement. For example, in Example 1.1, it tells us to prove choosing 51 numbers from 1 to 100 must give 2 consecutive integers. However, can this condition hold if 51 is replaced by a smaller number? Since questions are sometimes open, we have to seek for the best lower bound some time. How can we find the best lower bound? Pigeonhole Principle sometimes helps a lot.

To prove a number is the best lower bound, we have to check that any smaller integers lead to a counterexample. Let's take Example 1.1 as an example:

Example 3.1.

Prove that 51 cannot be replaced by any smaller integer in Example 1.1.

Solution.

From the solution in Example 1.1, we see that if any one pigeonhole contains 2 chosen integers, the condition is satisfied. Therefore, if the condition fails, each pigeonhole must contain at most 1 chosen number. We can use this clue to find a set of 50 numbers where each pigeonhole contains 1 chosen number only. It is not difficult to find out $\{1, 3, 5, 7, \dots, 99\}$ and $\{2, 4, 6, \dots, 100\}$ are sets of 50 integers without consecutive integers. Therefore, 51 is the best lower bound.

Example 3.2.

Prove that 6 cannot be replaced by 5 in Example 1.3.

Solution.

From the solution in Example 1.3, we see that if any point is incident to 3 blue edges or 3 red edges, then there is a monochromatic triangle. Therefore, no triangle occurs only if all 5 points have no 3 blue edges and no 3 red edges incident to it. Since there are 4 edges incident to each point, we can conclude for the edges incident to each point, exactly 2 of them are red and exactly 2 of them are blue. Having known this information, we can draw a counterexample easily. (Figure 3.1) Hence, 6 cannot be replaced by 5.

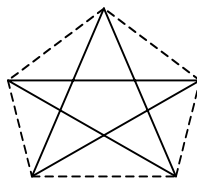


Figure 3.1: A counterexample. (Solid lines are red and dashed lines are blue.)

- It proves that $R(3,3) = 6$.
- Figure 3.1 is the only graph with 5 vertices in 2 colours having no monochromatic triangles. (To be precise in graph theory, we say it is unique up to homomorphism.) It consists of 2 monochromatic pentagons, one in red and one in blue.

Example 3.3.

If 19 is replaced by 18 in Example 2.3, find a colouring scheme for the 72 points such that no monochromatic rectangles exist.

Solution.

Using the notation of Example 2.3, if a pigeonhole has more than 1 pigeon, then there is a monochromatic rectangle. Therefore, the only way to make the condition fail is to let each pigeonhole contain no more than one pigeon. Since there are 18 columns (pigeons) and 18 pigeonholes, it forces each pigeonhole to accept one element. Moreover, if some of the columns have more than 1 choice of its pigeonholes, then clearly it can lead to more than 1 element in a pigeonhole. Therefore, each column must have only exactly 1 pair of points in the same colour. Using this clue, we can find a colouring scheme for 18 columns to avoid rectangles:

```

RRRBBB BBBRRR YYYBBB
RBBRRY BRRBBY YBBYR
BRYYR  RBYBYB BYRYR
YYRYRR YYBYBB RRYRY

```

Figure 3.2: A counterexample.

Example 3.4. (IMO 1992, Problem 3)

There are 9 points in space such that no four of them are coplanar. n edges are drawn between the points to form a simple graph ($n \leq 36$). The edges are either painted red or blue. Find the smallest possible value of n such that how the n edges are drawn and painted, there must exist a red or blue triangle.

Solution.

The smallest possible value of n is 33.

We claim that 33 edges work well. Since a complete graph with 9 vertices has at most $C_2^9 = 36$ edges, there will be 3 edges missing. For each missing edge, we remove one of its endpoint and also its connection with other vertices. This results in a complete subgraph with at least 6 vertices. As the edges of the subgraph are painted in red or blue, by Example 1.3, there is a red triangle or a blue triangle.

To prove the minimal property of 33, we have to construct a graph with 32 edges such that it has no red triangle or blue triangle. From the first part of the proof, we see that there should not be a complete subgraph of 6 vertices if no red or blue triangle appears. To avoid such a complete subgraph of 6 vertices, the 4 missing edges should have distinct vertices. Suppose the vertices are $A, B, C, D, a, b, c, d, X$, then the missing edges should be Aa, Bb, Cc, Dd (Aa means an edge connecting A and a).

Next, from Example 3.2, we see that each complete subgraph with 5 vertices must be like Figure 3.1. For example, $XABCD$ and $XaBCD$ are two of them. Without loss of generality, let a red pentagon in the former graph be $XABCD$. It implies that BC, CD and DX are blue. However, there is another red pentagon in the latter graph, so the only choice of the red pentagon is $XaBCD$. Therefore, Xa and aB are blue. Using similar arguments, we can claim that Ab, bC, Bc, cD, Cd, dX are blue. By considering the red pentagon in $XABCD$, we can know that $Xb, bD, Bd, dA, Da, aC, Ac$ and cX are red. Lastly, consider 5-vertex graphs like $XabCD, XaBcD, XaBCd, XAbcD, XAbCd$ and $XABcd$ similarly, we can check that ac, ad and bd are blue while ab, bc and cd are red.

To summarize, we have:

Blue edges: $XA, AB, BC, CD, DX, Xa, aB, Ab, bC, Bc, cD, Cd, dX, ac, ad, bd$.

Red edges: $XB, BD, DA, AC, CX, Xb, bD, Bd, dA, Da, aC, Ac, cX, ab, bc, cd$.

Missing edges: Aa, Bb, Cc, Dd .

We can check that this graph has no monochromatic triangle.

Despite the extensive application of Pigeonhole Principle in finding the best bounds, we should be aware that sometimes it fails. For example, if you want to follow the arguments of

Example 2.2 to find 24 points aligned in 3 rows and 8 columns with no monochromatic rectangles, you must fail as 7 columns already guarantee such a rectangle. The problem is that the arguments are not strong enough.

Exercises.

1. Prove that 27 can be replaced by 21 in Example 2.2. Moreover, draw a counterexample to show why 21 cannot be further replaced by 18.
2. Prove that 100 and 51 cannot be replaced by smaller numbers in Examples 2.4 and 2.5 respectively.
3. Find the smallest integer n such that no matter how n points are drawn on a plane. There are 3 points of them forming an angle not greater than 18° .
4. Can 18 be replaced by smaller numbers in Exercise 2 of Session 2? Justify your answer.
5. **(IMO 1991, Problem 3)** Let $S = \{1, 2, 3, \dots, 280\}$. Find the smallest integer n such that each n -element subset of S contains five numbers which are pairwise relatively prime.
6. **(IMO 2001, Problem 3)** Twenty-one girls and twenty-one boys took part in a mathematical contest, where the following situations are observed:
 - a. Each contestant solved at most six problems.
 - b. For each girl and each boy, at least one problem was solved by both of them.

Prove that there was a problem that was solved by at least three girls and at least three boys.

4. Solutions to Selected Exercises

The simple but elegant statement

2. Hint: Construct pigeonholes according to the number of friends.
4. Hint: Use the results of Example 1.3.

How to construct pigeonholes?

2. Let pairs of subjects (A, B) be pigeonholes and if a student studies both subjects, then he puts a pigeon into the pigeonhole. Each student has $C_2^4 = 6$ pigeons and there are $C_2^{15} = 105$ pigeonholes. There are totally $18 \times 6 = 108$ pigeons, so 2 of them are in the same pigeonhole.

4. Hint: Assuming the contrary, let the total number of bottles of milk drunk after k days be A_k . Then $A_1, A_2, \dots, A_{365}, A_1 + 29, A_2 + 29, \dots, A_{365} + 29$ must be pairwise distinct.
5. Hint: Align the 28 numbers in 4 rows and use the fact $C_2^4 = 6 < 7$.

Finding the bound

1. Hint: Use an argument similar to that in Example 2.3.
3. Hint: We may always assume that there is a straight line passing through one of the points and all other points are on the same side of the straight line. In order to prevent some of the angles to be too small, how should the points be arranged?
5. The answer is $n = 217$. To prove $n > 216$, Inclusion-Exclusion Principle is used. To prove 217 integers are enough, we construct pigeonholes using some subsets containing pairwise relatively prime integers.