Basic Combinatorics

Xiaoyi Cui

April 15, 2018

This is just a brief summary of the contents of the lectures. Please note: most of the calculations and demonstrations are neglected. I claim no originality of those contents.

Contents

Combinatorics involve three type of problems: the existence, the enumeration, and the optimization.

Combinatorial mathematics cuts across the many subdivisions of mathematics, and this makes a formal definition difficult. But by and large it is concerned with the study of the arrangement of elements into sets. The elements are usually finite in number, and the arrangement is restricted by certain boundary conditions imposed by the particular problem under investigation.

Enumeration and Stirling numbers

Definition 1 *A partition of an integer, n, is a representation of n as a sum of positive integers, where the order of the summands is not important. The number of partitions of n is denoted by $p(n)$. The number of partitions of n that contain exactly k summands is denoted by $p(n,k)$*[1].

[1] The value of $p(n,k)$ has other interpretations. Suppose we have n identical red balls and k identical buckets. In how many ways can the balls be placed into the buckets if every bucket must receive at least one ball? The answer turns out to be $p(n,k)$. This is an example of an occupancy problem.

Proposition 1 *Let n be a positive integer. Then*

$$p(n) = \sum_{k=1}^{n} p(n,k).$$

Theorem 1 (Generating function.) *Let n be a positive integer. Then $p(n)$ is the coefficient of z's in the generating function $\prod_{m=1}^{\infty}(\sum_{i=1}^{\infty} z^{im})$. That is,*

$$\sum_{n=1}^{\infty} p(n)z^n = \prod_{m=1}^{\infty}(1 + z^m + z^{2m} + z^{3m} + z^{4m} + \cdots)$$

Exercise 1 *Determine $p(3)$ from the above theorem.*

Corollary 1 *Let n be a positive integer. Then $p(n)$ is the coefficient of z^n in the polynomial*

$$\prod_{m=1}^{n}\left(\sum_{i=0}^{\lfloor \frac{n}{m} \rfloor} z^{im}\right).$$

Exercise 2 *Determine $p(4)$ from the above result.*

Theorem 2 (Recursive Relation) *Let n and k be integers and let $0 < k \leq n$. Then $p(n,k) = p(n-1,k-1) + p(n-k,k)$, where $p(n,k) = 0$ for $k > n$ and $p(n,0) = 0$. In addition, $p(n,n) = 1$.*

Definition 2 *The number of ways to distribute n distinguishable objects into k indistinguishable containers with every container receiving at least one object is denoted $S(n,k)$. The numbers, $S(n,k)$, are called the Stirling numbers of the second kind.*

Theorem 3 (Recursive relation) *Let n and k be integers and let $0 < k \leq n$. Then $S(n,k) = S(n-1,k-1) + kS(n-1,k)$. where $S(n,k) = 0$ for $k > n$ and $S(n,1) = 1$. In addition, $S(n,n) = 1$ and $S(n,0) = 0$ for $n > 0$.*

Definition 3 *The falling factorial is denoted $(x)_n$, and is defined by $(x)_0 = 1$ and*

$$(x)_n = \prod_{i=0}^{n}(x-i) = x(x-1)(x-2)\cdots(x-n+1)$$

for $n \geq 1$.

Proposition 2 *Let n be a positive integer. Then $x^n = \sum_{k=1}^{n} S(n,k)\cdot(x)_k$.*

Definition 4 *The coefficients of the expansion of $(x)_n$, as a linear combination of powers of x are the Stirling numbers of the first kind and are denoted by $s(n,k)$. That is, $(x)_n = \sum_{k=0}^{n} s(n,k)x^k$*[2]*.*

[2] The Stirling numbers of the first kind have no direct significance in counting problems: some of them are negative integers.

Theorem 4 *Let n and k be positive integers. Then $s(n,k) = s(n-1,k-1) - (n-1)s(n-1,k)$, where $s(0,0) = 1$, $s(n,0) = 0$, and $s(n,k) = 0$ if $n < k$. In addition, $s(n,n) = 1$.*

Combinatorial designes: Latin Square and finite projective planes

Definition 5 *A Latin square of order n is an n-by-n matrix for which every entry is a number in $\{1,2,\cdots,n\}$. Every number in $\{1,2,\cdots,n\}$ must appear at least once in every row and at least once in every column. A Latin square of order n is standardized if the elements in the first row are written in increasing numeric order moving left to right, and the elements in the first column are written in increasing numeric order moving top to bottom.*

Definition 6 *Let* $L^1 = (a_{ij})$ *and* $L^2 = (b_{ij})$ *be two Latin squares of order n.* L^1 *and* L^2 *are said to be orthogonal if the set of ordered pairs* $\{(a_{ij}, b_{ij})|i = 1, 2 \cdots nandj = 1, 2 \cdots, n\}$ *contains* n^2 *distinct ordered pairs. That is,* $(a_{ij}, b_{ij}) = (a_{rs}, b_{rs})$ *unless* $i = r$ *and* $j = s$*. A collection of k Latin squares of order n is said to be mutually orthogonal if every pair in the collection is orthogonal.*

Lemma 1 *Let L be a Latin square of order n and let* $i, j \in \{1, 2 \cdots n\}$*. If every copy of i in L is changed to a j, and simultaneously, every j in L is changed to an i, then the resulting n-by-n matrix is still a Latin square.*

Lemma 2 *Let* L^1 *and* L^2 *be orthogonal Latin squares of order n and let* $i, j \in \{1, 2 \cdots, n\}$ *Form an n-by-n matrix* $L^{1'}$ *by changing every copy of i in L to a j, and simultaneously, changing every j in* L^1 *to an i. All other entries in* L^1 *are copied unchanged into* $L^{1'}$*. Then* $L^{1'}$ *and* L^2 *are also orthogonal Latin squares.*

Theorem 5 *Let* $\{L^1, L^2 \cdots L^k\}$ *be a collection of mutually orthogonal Latin squares of order* $n > 1$*. Then* $k < n$*.*

Projective Plane

Definition 7 *A finite projective plane consists of a finite set, P, of points and a finite set, L, of lines. Lines are finite sets of points. If* $L = \{p_1, p_2 \cdots, p_k\}$*, then point* p_i *is said to be on line L and L is said to contain the point* p_i*, for* $i = 1, 2 \cdots k$*. The following axioms characterize finite projective planes.*

FPP1: Any two distinct points are on one and only one common line.

FPP2: Any two distinct lines contain one and only one common point.

FPP3: There exist four distinct points, no three of which are on a common line.

Axiom FPP2 implies that there is no such thing as "parallel lines" in a finite projective plane. Notice the symmetry in Axioms FPP1 and FPP2. It is possible to consider a point to be the set of lines that contain that point. This sense of "mirror image interchangeability" between points and lines is called duality.

Example 1 *The Fano Plane.*

Lemma 3 *Every line, L, in a finite projective plane must contain at least three distinct points.*

Lemma 4 *Let F be a finite projective plane.*

(A) Let L be a line in 5F that contains exactly $n + 1$ *points. If p is any point that is not on L, then p must be on exactly* $n + 1$ *distinct lines.*

(B) Let p be a point in F that is on exactly $n + 1$ *distinct lines. If L is any line that does not contain p, then L contains exactly* $n + 1$ *distinct points.*

Theorem 6 *Let F be a finite projective plane and let $n > 2$ be an integer. The following are equivalent.*

1. *There exists a line in F that contains exactly $n+1$ points.*
2. *There exists a point in F that is on exactly $n+1$ lines.*
3. *Every line in F contains exactly $n+1$ points.*
4. *Every point in F is on exactly $n+1$ lines.*
5. *There are exactly n^2+n+1 distinct points in F.*
6. *There are exactly n^2+n+1 distinct lines in F.*

Definition 8 *A finite projective plane, F, is said to have (or be of) order n if every line in F contains $n+1$ points.*

Example 2 *create a finite projective plane of order 3 by starting with a pair of mutually orthogonal Latin squares of order 3.*

Exercise 3 *Use the above example as an intuitive case, construct a finite projective plane of order $n > 1$ from $n-1$ mutually orthogonal latin squares of order n.*

Theorem 7 *If a set of $n-1$ mutually orthogonal Latin squares of order n exists, then a finite projective plane of order n also exists.*

Theorem 8 *If a finite projective plane of order n exists, then a set of $n-1$ mutually orthogonal Latin squares of order n also exists.*

Combinatorial optimization theory: Knapsack problems

Combinatorial optimization problems seek a best solution from among many contending solutions to some problem. The knapsack problem seeks to pack a knapsack with a set of items. Each item has a benefit value (or utility or cost) and a size (or weight). The goal is to obtain the maximum benefit under the constraint that the knapsack has a finite capacity.

Definition 9 *The knapsack problem is concerned with a knapsack that has positive integer volume (or capacity) v. There are n distinct items that may potentially be placed into the knapsack. Item i has positive integer volume v_i and positive integer benefit b_i. In addition, there are q_i copies of item i available, where quantity q_i is a positive integer satisfying $1 \leq q_i \leq \infty$. The integer variables $x_1, x_2 \cdots x_n$ will determine how many copies of item i are to be placed into the knapsack. The goal is to Maximize $L = \sum_i b_i x_i$ Subject to the constraints $\sum_i v_i x_i \leq v$ and $0 \leq x_i \leq q_i$. If $q_i = 1$ for $i = 1, 2, \cdots, n$, the problem is a 0-1 knapsack problem. If one or more of the q_i is infinite, the problem is unbounded; otherwise, the problem is bounded.*

Example 3 *01-Knapsack Problem.*

Exercise 4 *Use greedy algorithm to solve the above example problem. Greedy algorithm: Pack the item of highest utility first, then eliminate all items that won't fit. Now pack the remaining item of highest utility and eliminate any that now won't fit. Continue until no other item will fit.*

Exercise 5 *The greedy algorithm failed. Perhaps it was because we were focusing on the wrong property. A more sophisticated approach is to consider "benefit per unit of volume." That is, consider the ratios* $\frac{b_i}{v_i}$. *Try to solve the problem using this improved greedy algorithm.*

Exercise 6 *It is time to present an algorithm that correctly solves the knapsack problem. It will be easier to begin with a solution for completely unbounded knapsack problems (*$q_i = \infty$ *for* $i = 1, 2 \cdots n$*). A solution to this problem will then be modified to solve bounded and 0-1 knapsack problems. The key insight introduces recursion. Suppose an optimal solution is sought for a knapsack with capacity* v*, with the additional requirement that at least one copy of item* x *must be included. Suppose that item* x *has volume* v*, and benefit* b*. Let the optimal total benefit for a knapsack with capacity* v *be denoted by* $B(v)$*. A bit of thought will lead to the observation that the optimal solution will have total benefit* $b + B(v - v_x)$*. That is, add an item* x*, and then find the optimal benefit for the remaining* $v - v_x$ *cubic units of knapsack.*

The heuristic reduction technique

Theorem 9 *Suppose a knapsack has volume* v *and there are* n *types of items that can be packed. Denote the item benefits by* $\{b_l, b_2 \cdots b_n\}$*, the item volumes by* $\{v_l, v_2, \cdots, v_n\}$*, and the item quantities by* $\{q_1, q_2 \cdots, q_n\}$*. Assume that the items have already been ordered so that*

$$\frac{b_1}{v_1} \geq \frac{b_2}{v_2} \geq \cdots \frac{b_n}{v_n}.$$

If $b_1 \lfloor \frac{v}{v_1} \geq b_2 (\frac{v}{v_2})$*, then it is possible to find an optimal packing that includes min* $\{q_1, \lfloor \frac{v}{v_1} \rfloor\}$ *copies of item 1.*