Discrete mathematics

2017 - 2018, Fall

Instructor: Prof. János Pach

- covered material -

Lecture 1. Counting problems

To read:
[Lov]: 1.2. Sets, 1.3. Number of subsets, 1.5. Sequences, 1.6. Permutations, 1.7. Number of The Number of Ordered Subsets, 1.8. The Number of Subsets of a Given Size, 3.1. The Binomial Theorem, 3.2. Distributing Presents, 3.5. Pascal's Triangle, 3.6. Identities in Pascal's Triangle. **[Mat]**, Chapters 3.1-3.3.

Denote by $[n]$ the set of first n natural numbers: $[n] := \{1, 2, \ldots, n\}$.

Recall the following formulas:
- the number of functions from $[m]$ to $[n]$ is n^m. This is the number of m-letter words in an n-letter alphabet.
- the number of permutations of a set of n elements is $n!$
- the number of ways in which one can choose k objects out of n distinct objects, assuming the order of the elements matters, is $\frac{n!}{(n-k)!}$.
- the number of ways in which one can choose k objects out of n distinct objects, assuming the order of the elements does not matter, is $\frac{n!}{(n-k)!k!} = \binom{n}{k}$. This is the same as the number of subsets of k elements of an n-element set.

The following is called Pascal's triangle

Row										
0					$\binom{0}{0}=1$					
1					$\binom{1}{0}=1$		$\binom{1}{1}=1$			
2				$\binom{2}{0}=1$		$\binom{2}{1}=2$		$\binom{2}{2}=1$		
3			$\binom{3}{0}=1$		$\binom{3}{1}=3$		$\binom{3}{2}=3$		$\binom{3}{3}=1$	
4		$\binom{4}{0}=1$		$\binom{4}{1}=4$		$\binom{4}{2}=6$		$\binom{4}{3}=4$		$\binom{4}{4}=1$
5	$\binom{5}{0}=1$		$\binom{5}{1}=5$		$\binom{5}{2}=10$		$\binom{5}{3}=10$		$\binom{5}{4}=5$		$\binom{5}{5}=1$

The following identities hold:
1. $\binom{n}{k}$ is the k-th element in the n-th line of Pascal's triangle.
2. $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.

3. All the diagonals of Pascal's triangle are strictly increasing.

The number of subsets of an n-element set is 2^n, since we have

$$2^n = \binom{n}{0} + \binom{n}{1} + ... + \binom{n}{n}.$$

The number of subsets of an n-element set having odd cardinality is 2^{n-1}. The number of subsets of an n-element set having even cardinality is 2^{n-1}.

The equalities above can be obtained using the **binomial theorem**.

$$(1+x)^n = \binom{n}{0} + \binom{n}{0}x + ... + \binom{n}{n}x^n = \sum_{i=0}^{n}\binom{n}{i}x^i.$$

For $x = 1$, respectively $x = -1$, we obtain

$$2^n = \binom{n}{0} + \binom{n}{1} + ... + \binom{n}{n} = \sum_{i=0}^{n}\binom{n}{i},$$

$$0 = \binom{n}{0} - \binom{n}{1} + ... + (-1)^n\binom{n}{n} = \sum_{i=0}^{n}(-1)^i\binom{n}{i},$$

Adding, respectively substracting the two relations, and dividing each by two, one obtains

$$2^{n-1} = \binom{n}{0} + \binom{n}{2} + ...$$

$$2^{n-1} = \binom{n}{1} + \binom{n}{3} + ...,$$

which proves the statements about the number of even/odd sets.

Assume we have k identical objects and n different persons. Then, the number of ways in which one can distribute this k objects among the n persons equals

$$\binom{n+k-1}{n-1} = \binom{n+k-1}{k}.$$

Equivalently, it is a number of solutions of the equation $x_1+...+x_n = k$ in nonnegative integers or the number of k-multisets containing elements from $[n]$.

If $k \geq n$ and each persons receives at least 1 object, then the number of possible ways to distribute is $\binom{k-1}{n-1}$.

Theorem 1 (Multinomial theorem)**.** *The following holds:*

$$(x_1 + ... + x_n)^k = \sum_{\substack{i_1,i_2,...,i_n \geq 0 \\ i_1+...+i_n=k}} \frac{k!}{i_1!...i_n!}x_1^{i_1}...x_n^{i_n}.$$

Lecture 2. "big oh" and "little oh" notation, effective estimates for $n!$, $\binom{n}{k}$, etc., Stirling formula

To read:
[**Lov**] 2.1. Induction, 2.2. Comparing and estimating umbers, 2.4. Pigeonhole principle
[**Mat**] 3.4. Estimates: an introduction - starting from 3.4.2. - Big Oh, little oh, 3.5.5. Estimate $n!$ - second proof only,

Definition 2. Let $f, g : \mathbb{Z}_+ \to \mathbb{R}$. We say that f is *big-Oh* of g and we write $f(x) = O(g(x))$ if there exist n_0 and c constants such that for all $n > n_0$, we have $|f(n)| < c \cdot |g(n)|$.

Definition 3. Let $f, g : \mathbb{Z}_+ \to \mathbb{R}$. We say that f is *little-oh* of g and we write $f(x) = o(g(x))$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0.$$

Examples:

- If $f(n) = o(g(n))$ then $f(n) = O(g(n))$.
- If $f(n) = O(g(n))$ and $g(n) = O(h(n))$ then $f(n) = O(h(n))$ (and if $g(n) = o(h(n))$ then $f(n) = o(h(n))$)
- $n + 1 = o(n^2)$ but n^2 is not $o(n + 1)$
- $n^2 + 2 = o(2^n)$, $n^4 + 5n^3 - 2n - 10 = o(2^n)$

Estimating $n!$
Easy observation: for all $n \geq 1$ we have

$$2^n \leq n! \leq n^{n-1}.$$

Improved bounds:

$$e\left(\frac{n}{e}\right)^n \leq n! \leq en\left(\frac{n}{e}\right)^n$$

This proof uses estimates of $\ln 1 + \ldots \ln n$ using integrals.

Theorem 4 (Stirling's formula)**.**

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$

where $\sim$ is used to indicate that the ratio of the two sides tends to 1 as n goes to ∞.

Binomial coefficients:

- can be estimated using Stirling's formula
- or the binomial theorem: $\binom{n}{k} < \left(\frac{ne}{k}\right)^k$

Lecture 3. The inclusion-exclusion principle, permutations without fixed points, Euler's function $\varphi(n)$, introduction to graph theory

Number of permutations without fixed points
To read: **[Lov]** 2.3. Inclusion-Exclusion
[Mat] 3.7. Inclusion - Exclusion. 3.8. The hatcheck lady.

For two finite sets A_1 and A_2, we have $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$.
For three finite sets A_1, A_2, A_3, we have

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$

This works more generally:

Theorem 5 (Inclusion-Exclusion)**.** *Let $A_1, ..., A_n$ be finite sets. Then, the following holds*

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{1\le i<j\le n} |A_i \cap A_j| + \sum_{1\le i<j<k\le n} |A_i \cap A_j \cap A_k| - ... + (-1)^{n-1} |A_1 \cap A_2 \cap ... \cap A_n|.$$

The Hatcheck Lady problem
n people receive their hats randomly permuted. What is the probability that no one gets back his own hat?

Let us look at the permutations of the set $\{1, 2, ..., n\}$ without fixed points. In order to count these, we apply the inclusion-exclusion principle. Let A be the set of all permutations and A_i be the set of permutations of the set $\{1, 2, ..., n\}$ for which i is a fixed point. The number of permutations with no fixed points is

$$\left|A \setminus \bigcup_{i=1}^{n} A_i\right| = |A| - \left|\bigcup_{i=1}^{n} A_i\right|.$$

We know that $|A| = n!$, so we need to count $|\bigcup_{i=1}^{n} A_i|$. We do this using the inclusion principle.

Note that $A_i \cap A_j$ represents the set of all permutations for which i and j are fixed points. One can see that $|A_i| = (n-1)!$ for all i, while $|A_i \cap A_j| = (n-2)!$. Using the same idea, we obtain $|A_i \cap A_j \cap A_k| = (n-3)!$ and so on. Altogether, this gives

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} (n-1)! - \sum_{1\le i<j\le n} (n-2)! + \sum_{1\le i<j<k\le n} (n-3)! - ... =$$

$$= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)! - ... = n!\left(\sum_{k=0}^{n} (-1)^{k+1} \frac{1}{k!}\right) \sim n!\left(1 - \frac{1}{e}\right).$$

Thus, the number of permutations without fixed points is $\sim n! - n!(1 - \frac{1}{e}) = n!/e$, as $n \to \infty$.

To get the desired probability, we need to divide by the total number of permutations, $n!$. Then we see that the probability tends to $1/e$.

Euler's function

For every positive integer n we define $\phi(n)$ as the number of positive integers that are relatively prime with n. Formally, one writes

$$\phi(n) = \{m \in \{1, 2, ..., n\} | gcd(m, n) = 1\},$$

where $gcd(m, n)$ denotes the greatest common divisor of m and n.

If $n = p^\alpha$, with p prime, we have

$$\phi(n) = \phi(p^\alpha) = p^\alpha - p^{\alpha-1}.$$

Otherwise, $n = \prod_{i=1}^{m} p_i^{\alpha_i}$ with p_i prime factors. In this case, using inclusion-exclusion, one obtains that

$$q(n) = n \cdot \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right).$$

Elements of graph theory

To read:

[**Lov**] 8.1. How to Define Trees?, [**Mat**] 4.1. The notion of a graph; isomorphism - only the definition of graphs, 4.3.1. Sum of the degrees, 5.1. Definition and characterizations of trees.

Definition 6. A *graph* G is an ordered pair (V, E), where V is a set of elements called *vertices* and E is a set of 2-element subsets of V called *edges*.

Definition 7. Let $G = (V, E)$ be a graph. We call a sequence of vertices $v_0, ..., v_r$ *a walk* if (v_i, v_{i+1}) is an edge of G, for every $0 \leq i \leq r - 1$. A *simple walk* or a *path* is a walk without any repeated vertices.

Definition 8. A graph H is a *subgraph* of $G = (V, E)$ (denoted $H \subseteq G$) if H can be obtained from G by deleting vertices and edges.

Definition 9. We say that a graph $G = (V, E)$ is *connected* if for every two vertices $v, u \in V$ there exists a walk in G between u and v. Equivalently, G is connected if for every two vertices $v, u \in V$, there exists a path in G between u and v.

Definition 10. For every vertex of a graph, we define its *degree* as the number of edges adjacent to it.

The following lemma gives us a relation between the degrees of the vertices of a graph and the total number of edges:

Lemma 11. *For every graph $G = (V, E)$, the following holds:*

$$2|E| = \sum_{v \in V} d(v),$$

where $d(v)$ represents the degree of the vertex v.

Proof. The proof is based on a double counting of the total number of edges. On the one hand, we know that this is $|E|$. On the other hand, when counting the sum of degrees of all vertices, we count every edge twice. □

Definition 12. A *cycle* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, ..., v_r \in V$ such that $v_r = v_0$ and $\{v_i, v_{i+1}\} \in E$ for all i from 0 to $r - 1$.

Definition 13. A *tree* is a connected graph without cycles.

Lecture 4. Equivalent definitions of a tree, number of labeled trees via Prufer codes

Equivalent definitions of a tree To read: **[Lov]** 8.1. How to Define Trees, 8.2. How to Grow Trees
[Mat] 5.1. Definition and characterizations of trees.

Theorem 14. *The following properties are equivalent:*
(1) T is a tree.
(2) Any two vertices in T are connected by a unique path.
(3) T is minimally connected, that is, it is connected, but after deleting any edge, it becomes disconnected
(4) T is maximally acyclic, that is, it is acyclic, but if we add any edge to T, then it will contain a cycle.
(5) T has one edge less than the number of vertices and it is connected.
(6) T has one edge less than the number of vertices and it is acyclic.

Definition 15. A vertex of degree one in a tree is called a *leaf*.

Lemma 16. *Every tree on $n \geq 2$ vertices has at least one leaf.*

Proof. Let S be the set of all the paths in the tree T. We know that every path on r vertices contains exactly $r-1$ edges.

Consider now a path $v_1, ..., v_l$ of maximum length. One can always find a path of maximum length since every path in the tree can contain at most n vertices (otherwise it will be self-intersecting, that is it will contain a cycle, which is impossible since in a tree we cannot have cycles).

We prove that both v_1 and v_l (the endpoints of the path) are leafs. Assume at least one of them is not, say v_1. That means that, there is at least another edge apart from v_1v_2 incident to v_1. Observe that u cannot coincide with any of the vertices of the path $v_1, ..., v_l$ (otherwise it will close a cycle). Therefore, we can add u to the path without forming any cycle. But this is a contradiction to the maximality of the length of the path $v_1, ..., v_l$. Thus, both v_1 and v_l must be leaves. □

Lemma 17. *If T is a tree and v is a leaf ot T then $T - v$ is also a tree. Here $T - v$ denotes the graph obtained from T by deleting v and the edge incident to it.*

Theorem 18. *Every tree on n vertices has exactly $n-1$ edges.*

Proof. We do induction on the number of vertices. The statement holds for $n = 1, 2$ so assume that every graph on n vertices contains $n-1$ edges.

We want to prove that the statement is true for $n+1$, that is, every tree on $n+1$ vertices has exactly n edges. Let $T_{n+1} = (V, E)$ be an arbitrary tree on $n+1$ vertices. By the first lemma above, we know that T contains at least one leaf v. We remove v and its unique incident edges from the tree T_{n+1}. By the second lemma, we know that this leaves us with a tree T_n on n vertices, which, by the induction hypothesis has exactly $n-1$ edges. Therefore, T_{n+1} contains exactly $n-1+1 = n$ edges, which completes the proof. □

Counting labeled trees

In what follows, we will present a result due to Cayley. Before stating the theorem, we need the following lemma:

Lemma 19. *Let T be a tree on n labeled vertices and let $d_1, ..., d_n$ be the degrees of the vertices. Then*

$$\sum_{i=1}^{n} d_i = 2|E(T)| = 2(n-1),$$

where by $E(T)$ denotes the edge set of the tree.

Now we can state Cayley's theorem.

Theorem 20 (Cayley)**.** *The number of trees on n labeled vertices is n^{n-2}.*

We give two proofs to this theorem. The first one, due to Prüfer, is algorithmic.

Proof 1 of Cayley's theorem. We give now the proof, due to Prüfer.

Denote the vertices by $\{1, 2, ..., n\}$. We will define a one-to-one correspondence between the set of all trees on n labeled vertices and the set of all sequences of length $n-2$ consisting of numbers in $\{1, 2, ..., n\}$. Since the cardinality of the latter is n^{n-2}, we obtain the desired result.

The following algorithm takes a tree as input, and yields a sequence of integers:

Step 1: Find the leaf with the smallest label and write down the number of its neighbor.

Step 2: Delete this leaf, together with the only edge adjacent to it.

Step 3: Repeat until we are left with only two vertices.

Claim 21. *The labels not occurring in the sequence are exactly the leaves of the tree. Moreover, a vertex of degree d occurs exactly $d-1$ times in the sequence.*

We present an algorithm that reconstructs the tree from the Prüfer code.

Step 1: Draw the n nodes, and label them from 1 to n.

Step 2: Make a list of all the integers $(1, 2, ..., n)$. This will be called the list.

Step 3: If there are two numbers left in the list, connect them with an edge and then stop. Otherwise, continue on to step 4.

Step 4: Find the smallest number in the list which is not in the sequence. Take the first number in the sequence. Add an edge connecting the nodes whose labels correspond to those numbers.

Step 5: Delete the smallest number from the list and the first number in the sequence. This gives a smaller list and a shorter sequence. Then return to step 3.

[**Lov**] 8.3. How to Count trees? 8.4. How to Store trees?

[**Mat**] 5.1 Definition and characterizations of trees 8.1. The number of spanning trees, 8.4. A proof using the Prüfer codes.

Lecture 5. Another proof of Cayley's formula, unlabeled trees, Kruskal algorithm for finding a spanning tree

Counting trees revisited

Recall Cayley's formula:

Theorem 22 (Cayley). *The number of trees on n labeled vertices is n^{n-2}.*

We give another proof of it. It uses induction.

Proof 2 of Cayley's theorem. The proof is based on the following lemma

Lemma 23. *The number of trees on n vertices labeled with $1, 2, ..., n$ with degrees $d_1, ..., d_n$ equals*

$$\frac{(n-2)!}{(d_1-1)!...(d_n-1)!}.$$

[Two proofs for the lemma: one using Prüfer codes and another using induction.]

By the lemma, we obtain that the total number of trees on n labeled vertices is the sum of the number of trees over all possible values of degrees, that is

$$\sum_{\substack{d_1,...,d_n \geq 1 \\ d_1+...+d_n=2n-2}} \frac{(n-2)!}{(d_1-1)!...(d_n-1)!}.$$

Let now $r_i = d_i - 1$, for all $1 \leq i \leq n$. By substitution, we obtain that the number of trees on n labeled vertices is

$$\sum_{\substack{r_1,...,r_n \geq 1 \\ r_1+...+r_n=n-2}} \frac{(n-2)!}{(r_1)!...(r_n)!},$$

which, by the multinomial theorem equals n^{n-2}.

Counting unlabeled trees

The number of unlabeled trees, that is, classes of pairwise nonisomorphic trees is only exponential in the number of vertices. We prove the following theorem:

Theorem 24. *The number of pairwise nonisomorphic trees on n vertices is at most 2^{2n-4}.*

Here's a sketch of a proof:

The proof uses the following encoding of trees. We think of a tree hanged from one of its vertices on a plane (we think of gravity working in the negative y-direction). We go around the tree and form a binary sequence. If we are going one edge down, we write 1 in the sequence. If we're going up - we write 0. At the end we corresponded one 0 and one 1 to each edge, which gives us a binary sequence of length $2n-2$. The last bit is always

0, and the first bit is always 1, so the total number of these sequences is at most 2^{2n-4}. It is not difficult to see that any such sequence determines a tree.

Weighted spanning trees

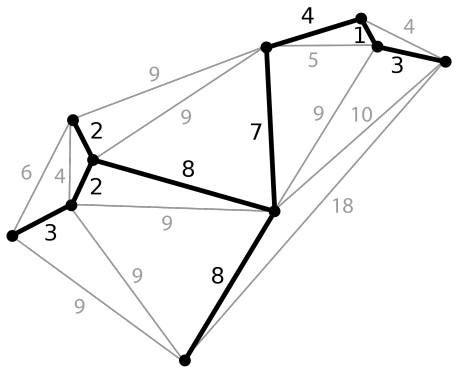
Definition 25. A *weighted graph* is a graph in which each edge is given a numerical weight.

We define the weight of a graph as the sum of the weights of all its edges.

We are interested in the following problem: find a minimum weight spanning tree T for a given weighted connected graph G. Before that, we have to define the notion of a spanning tree for a graph G.

Definition 26. Let $G = (V, E)$ be a graph. We say that a tree T is a *spanning tree* of G if it contains all the vertices of V and is a subgraph of G, that is every edge in the tree belongs to the graph G.

Example. Below is an example of a spanning tree:



One way to solve the problem of finding a minimum spanning tree is using Kruskal's algorithm. This works as follows:

Step 1. Start with an empty graph.

Step 2. Take all the edges that haven't been selected and that would not create a cycle with the already selected edges and select it unless it creates a cycle. Add the one with the smallest weight.

Step 3. Repeat until the graph is connected.

Refer as well to the following:

[**Lov**] 8.3. How to Count trees?

[**Mat**] 5.2. Isomorphism of trees, 5.3. Spanning trees of a graph, 5.4. Minimum spanning tree problem, 8.1. The number of spanning trees, 8.2. A proof via scores.

Lecture 6. Partial orders, Dilworth's theorem

Definition 27. A *partially ordered set* (or simply poset) is a pair $(X, \leq)$, where X is a set and $\leq$ is a binary relation over X, which is reflexive, antisymmetric, and transitive, i.e., which satisfies the following relations, for all a, b, and c in X:

a. $a \leq a$ (reflexivity);

b. if $a \leq b$ and $b \leq a$ then $a = b$ (antisymmetry);

c. if $a \leq b$ and $b \leq c$ then $a \leq c$ (transitivity).

We write $a < b$ if $a \leq b$ and $a \neq b$.

Definition 28. Let $(X, \leq)$ be a partially ordered set.

A *chain* in X is a sequence $x_1, ..., x_t \in X$ such that

$$x_1 < x_2 < ... < x_t.$$

An *antichain* in X is a subset $\{x_1, ..., x_t\}$ of X such that no two elements are comparable.

Recall, that for a set X, 2^X denotes the power set of X, that is the set of all subsets of X.

Example. Consider the partially ordered set $(2^{\{1,2,3\}}, \subseteq)$.

The sequence $\emptyset \subset \{1\} \subset \{1,2,3\}$ is a chain.

$\{1,2\}, \{1,3\}$ is an antichain.

Theorem 29 (Dilworth)**.** *Let $(X, \leq)$ be a partially ordered set.*

a. If the maximum size of an antichain is k, then X can be decomposed into k chains.

b. If the maximum size of a chain is k, then X can be decomposed into k antichains.

Proof. The proof can be found in [Juk], page 108. We present here the proof of the part a, which is more difficult.

We use induction on the cardinality of X. let a be a maximal element of X, and let k be the size of a largest antichain in $X' = X \setminus \{a\}$. Then X' is the union of k disjoint chains $C_1, \ldots, C_k$. We have to show that X either contains an $(k+1)$-element antichain or else is the union of k disjoint chains. Now, every k-element antichain in X' consists of one element from each C_i. Let a_i be the maximal element in C_i which belongs to some k-element antichain in X'. It is easy to see that $A = \{a_1, \ldots, a_k\}$ is an antichain in X'. if $A \cup \{a\}$ is an antichain in X, we are done: we have found an antichain of size $k+1$. Otherwise, we have $a > a_i$ for some i. Then $K = \{a\} \cup \{x \in C_i : x \leq a_i\}$ is a chain in X, and there are no k-element antichains in $X \setminus K$ (since A_i was the maximal element of C_i participating in such an antichain), whence $X \setminus K$ is the union of $k-1$ chains. □

Corollary 30. *Let $(X, \leq)$ be a partially ordered set. Then, the following hold:*

a. The maximum size of an chain is equal to the minimum number of antichains that cover X.

b. The maximum size of an antichain is equal to the minimum number of chains that cover X.

Refer as well to the following:
[**Mat**] 2. Orderings: the definitions of poset, chain, antichain, Hasse diagram.
[**Juk**] 8.1. Decomposition in chains and antichains.

Lecture 7. König-Hall theorem, Sperner's theorem

Definition 31. A *bipartite graph* is a graph G whose vertices can be divided into two disjoint sets A and B such that every edge of the graph connects a vertex in A to one in B (in other words, there is no edge of the graph between two vertices of A or two vertices of B.

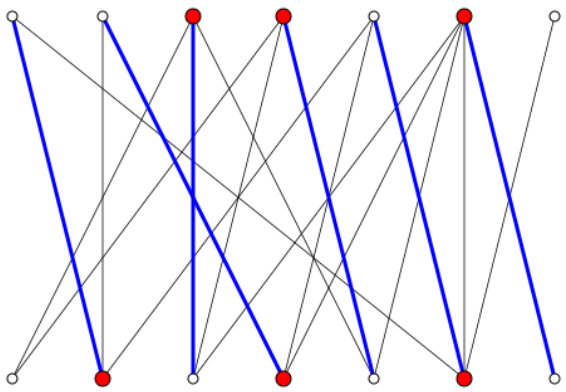
Lemma 32. *A graph is bipartite if, and only if, it does not contain an odd cycle (that is, a cycle of odd length).*

Definition 33. Let $G = (V, E)$ be a graph. A subset $E' \subseteq E$ of pairwise disjoint edges (that is, edges which do not share any vertex) is called a *matching* in G.

Definition 34. A *perfect matching* is a matching where every vertex of the graph is incident to exactly one edge of the matching.

Remark 35. A perfect matching is therefore a matching of a graph containing $n/2$ edges (where n is the number of vertices). Thus, perfect matchings are only possible on graphs with an even number of vertices!

Example 36. The blue edges in the following graph form a matching: The matching is



not a perfect matching though.

What obstacles prevent the existence of a perfect matching in a bipartite graph $G(A \cup B, E)$?

Example 37. If a vertex set X in A has fewer than $|X|$ neighbors in B then X cannot be covered by any matching in G. In particular, if such a set X exists then G has no perfect matching.

Interestingly, the converse is also true:

Theorem 38 (König-Hall)**.** *Let $G = (V, E)$ be a bipartite graph with bipartition $V = A \cup B$. For every $X \subseteq A$, let*

$$B(X) = \{b \in B : \exists x \in X \text{ such that } (x, b) \in E\}.$$

Then there is a matching M in G such that every vertex of A belongs to some edge in M if and only if $|B(X)| \geq |X|$, for all $X \subseteq A$.

Proof. We will deduce this from Dilworth's theorem. Define the poset $(P, \prec)$ where $P = V$ is the vertex set of the graph and $u \prec v$ if $u \in A, v \in B$ and $uv \in E$. By Dilworth, we know that the maximum size of an antichain is equal to the maximum number of chains that P can be partitioned into.

Claim 39. *The maximum size of an antichain is $|B|$.*

Proof. Let D be an antichain in P. Then we know that $X = D \cap A$ satisfies the condition (Hall's condition) above, so $|B(X)| \geq |X|$. As D is an antichain, we also know that no vertex of $B(X)$ appears in D, and hence $B(X)$ and $D \cap B$ are disjoint. But then

$$|B| \geq |B(X)| + |D \cap B| \geq |X| + |D \cap B| = |D \cap A| + |D \cap B| = |D|$$

which is what we wanted to show. □

So by Dilworth' theorem, we know that P can be partitioned into $|B|$ chains. The following observation is easy to see from the definition of our poset:

Claim 40. *There is no chain longer than 2.*

Thus each chain is either an edge or a vertex.

Let us take a partition of the vertices into $|B|$ chains. We know that no chain can contain two vertices from B, so each vertex in B has a separate chain containing it. This means that at least $|B|$ (that is: all) of our chains contain a vertex from B. In particular, there is no chain that only contains a single vertex of A. But then the chains containing the vertices of A form a matching that we were looking for. □

Theorem 41 (Sperner)**.** *Let $X = \{1, 2, ..., n\}$ and $A_1, ..., A_m \subseteq X$, with $A_j \nsubseteq A_i$, for all $i \neq j$. Then $m \leq \binom{n}{\lfloor n/2 \rfloor}$.*

Proof. Define the poset on the power set 2^X of X (2^X contains all subsets of X), where $A \prec B$ is $A \subset B$. By Dilworth' theorem, it is enough to decompose this poset into $\binom{n}{\lfloor n/2 \rfloor}$ chains. In fact, we partition into symmetric chains, that is, maximal chains connecting level k and level $n - k$ for some k.

The construction is based on induction as follows: For $n = 0$ the poset has 1 element, so we can cover it by $\binom{0}{0} = 1$ symmetric chain. Now suppose we have a chain decomposition for some n. Then if for each of our chains of the form $A_k \subset A_{k+1} \subset \cdots \subset A_{n-k}$ we add the chain $A_k \cup \{n+1\} \subset A_{k+1} \cup \{n+1\} \subset \cdots \subset A_{n-k} \cup \{n+1\}$ to our partition, then we get a chain decomposition for $n + 1$.

However, these chains are not symmetric for $n + 1$. So instead we move A_k to the front of the other chain. Then $A_{k+1} \subset \cdots \subset A_{n-k}$ and $A_k \subset A_k \cup \{n+1\} \subset \cdots \subset A_{n-k} \cup \{n+1\}$ are symmetric, and they cover the same sets, so they form a symmetric chain decomposition, proving the theorem.

Note that some of our new chains can become empty and so are not really chains. This happens if n is even and $k = n/2$. □

Refer as well to the following:

[Lov] 10.3. The Main Theorem **[Mat]** 7.2. Sperner's theorem on independent systems: Sperner theorem and proof of Theorem 7.2.1.

Lecture 8. Other proofs of Sperner's theorem - LYM inequality - The Littlewood-Offord problem

Definition 42. Let X be a set.

- $2^X = \{Y : Y \subseteq X\}$ is the family of subsets of X. Note that $|2^X| = 2^{|X|}$.
- $\binom{X}{k} = \{Y : Y \subseteq X \text{ and } |Y| = k\}$ is the family of k-element subsets of X. Note that $|\binom{X}{k}| = \binom{|X|}{k}$.

We will give another proof using the following corollary of Hall's theorem.

Theorem 43. *Let $G = (A \cup B, E)$ be a bipartite graph such that all vertices in A have the same degree k, and all vertices in B have the same degree l, where $k \geq l$. Show that G has a matching M such that every vertex in A is contained in an edge of M.*

Proof. We need to check that Hall's condition holds for A. Take any subset $S \subseteq A$, let $B(S)$ be its neighborhood in B and look at the number of edges connecting S and $B(S)$. On the one hand, it is exactly $|S|k$ because every vertex in S has degree k. On the other hand, it is at most $|B(S)|l$ because every vertex in $B(S)$ has degree l (so at most l of them go to S). Hence $|S|k \leq |B(S)|l \leq |B(S)|k$ (using $l \leq k$, implying $|S| \leq |B(S)|$ for every S. □

Second proof of Theorem 41. If $k < n/2$ then consider the bipartite graph induced by the k'th and $k+1$'st levels of the Hasse diagram, i.e., we connect a k-element set Y and a $k+1$-element set Z if $Y \subset Z$. This graph satisfies the condition of the above corollary of Hall's theorem. Indeed, every k-element set is connected to $n-k$ of the $k+1$-element sets and every $k+1$-element set is connected to $k+1$ of the k-element sets. Thus the k-element sets can be "matched into" the $k+1$-element sets in a way that every edge of this matching corresponds to a chain (of length 2) in the poset. By symmetry, we can do the same for each $k > n/2$ and match the $k+1$'st level into the k'th level.

Now let us look at the union of the n matchings we obtained. It is a union of paths. Because it is a subgraph of the Hasse-diagram, every such path will be a chain in the poset. Also, because we matched smaller levels into larger levels, every path will contain an element from the middle level(s). Indeed, look at a vertex not in the middle level. It was matched to a larger level, so its path also contains a vertex closer to the middle level. And so on, we can walk from any vertex to the middle level on its path, hence every path reaches the middle level.

But then there are exactly $\binom{n}{\lfloor n/2 \rfloor}$ paths and they form a chain partition of the poset. The theorem follows. □

Definition 44. A *maximal chain* is a chain $A_0 \subset A_1 \subset \cdots \subset A_n$ where $|A_i| = i$.

The number of maximal chains is $n!$

Claim 45. *If $B_1, \ldots, B_m$ is a "Sperner family" (that is, $A_j \not\subseteq A_i$, for every $i \neq j$) in 2^X, then no chain passes through more than one B_i's.*

We state the LYM inequality and deduce another proof of Sperner's theorem.

Theorem 46 (LYM inequality)**.** *Let $X = \{1, 2, ..., n\}$, and let $\mathcal{F}$ be a family of subsets $A_1, ..., A_m \subseteq X$ such that $A_j \nsubseteq A_i$, for all $i \neq j$. Let $m_k = |\{A \in \mathcal{F} : |A| = k\}|$, that is the number of sets in $\mathcal{F}$ containing k elements.(One can see that $m_1 + ... + m_n = m$). Then, the following holds*

$$\sum_{i=0}^{n} \frac{m_i}{\binom{n}{i}} \leq 1.$$

Proof. The proof and how Sperner's theorem follows can be found in [Mat] pages 227-228. □

Also refer to the following:

[**Bol**] 3. Sperner systems.

[**Mat**] 7.2. Sperner's theorem on independent systems: Sperner theorem and proof of Theorem 7.2.1.

The Littlewood-Offord problem: Let $a_1, ..., a_n \geq 1$ be fixed real numbers and

$$A = \left\{ \sum_{i=1}^{n} \epsilon_i a_i,\ \epsilon_i = -1 \text{ or } 1 \right\}.$$

We are interested in finding how many of the 2^n sums of A lie inside a given interval I of length less than 2.

We will prove that, no matter how one chooses I, we have

$$|A \cap I| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

We do this using Sperner's theorem. We construct a bijection as follows: for every sum $\sum_{i=0}^{n} \epsilon_i a_i$ we assign a unique characteristic set, that is $\{i \mid \epsilon_i = 1\} \subseteq \{1, 2, ..., n\}$. One can easily see that this is a bijection, so it is enough to bound the number of characteristic sets.

We prove that the characteristic sets of sums contained in the same interval satisfy the condition in Sperner's theorem, that is no one is contained in another. We assume the contrary, that is, there exist two characteristic sets B, B' with $B \subseteq B'$, $B \neq B'$ such that the corresponding sums lie in the interval I

$$\sum_{i \in B} \epsilon_i a_i \in I \text{ and } \sum_{i \in B'} \epsilon_i a_i \in I.$$

Since the length of I is less than 2, their difference has to be less than two. On the other hand, since $B \subseteq B'$ and $B \neq B'$, we obtain that

$$\sum_{i \in B'} \epsilon_i a_i - \sum_{i \in B} \epsilon_i a_i = \sum_{i \in B' \setminus B} 2a_i < 2,$$

where the sum is a sum of non-negative numbers, containing at least one non-zero term. On the other hand, since every $a_i \geq 1$, we obtain that the sum must be at least two, which is a contradiction.

Therefore, the condition of Sperner's theorem is satisfied, so we obtain that the number of sums inside the interval I cannot exceed $\binom{n}{\lfloor n/2 \rfloor}$, which completes the proof.

Lecture 9. Stable matchings

Suppose that you have n students that want to do their internships in n companies. Each have sent their applications to all the companies. Each student and each company has their list of preferences, and we want to pair up the students with the companies so that this assignment is *stable* in the following sense: there is no student and company that would both prefer to work with each other to their assigned pairs.

More formally, consider a bipartite graph G with parts A, B, where $|A| = |B| = n$, and in which each vertex has a (strict) order of preferences for the vertices of the other part. We say that a perfect matching is *stable*, if there is no pair $a \in A, b \in B$, such that both of them would prefer the other to the vertex they are currently matched to.

Below we present an algorithm of Gale and Shapley, which allows to construct such a stable matching.

The Gale-Shapley Algorithm to find a stable matching M in a complete bipartite graph G with bipartition $V(G) = A \cup B, |A| = |B|$

(1) Set $M = \emptyset$;

(2) Iterate:

  (a) Take an unmatched vertex $a \in A$ and let $b \in B$ be the the vertex that a prefers among the ones a has not tried yet.

  (b) a "proposes" to b: If b is unmatched or b is matched to a', but prefers a over a', then "accept" a and "reject" a': put $M := M - a'b + ab$. Otherwise, "reject": leave M unchanged;

  (c) If there is no more unmatched vertices in A that have someone left on the list, then go to (3);

(3) Return M.

The complexity of this algorithm is $O(n^2)$.

Theorem 47 (Gale and Shapley)**.** *The matching M that the algorithm outputs is stable.*

Proof. First we show that M is perfect, i.e. that every vertex in $A \cup B$ is matched. Indeed, if there is a pair of vertices $a \in A, b \in B$, such that both are not in the matching, then a must have proposed to b at some point. However, if a vertex $b \in B$ is in M at some step of the algorithm, then it stays in M.

Next, we show that the matching is stable. Assume that $ab \notin M$. Upon completion of the algorithm, it is not possible for both a and b to prefer each other over their current match. If a prefers b to its match, then a must have proposed to b before its current match. If b accepted its proposal, but is matched to another vertex at the end, then b prefers the current match of b over a. If b rejected the proposal of a, then b was already matched to a vertex that is better for b. □

Lecture 10. Erdős-Ko-Rado theorem. The probabilistic method.

To read: [**Juk**] 7.2 Erdős-Ko-Rado theorem.

Definition 48. Let X be a set and $\mathcal{F}$ be a family of subsets of X, that is $\mathcal{F} \subseteq 2^X$. We say that $\mathcal{F}$ is *intersecting*, if any two members of $\mathcal{F}$ intersect.

Lemma 49. *Let X be a finite set with $|X| = n$ and $\mathcal{F} \subseteq 2^X$ an intersecting family. Then $|\mathcal{F}| \leq 2^{n-1}$ and this bound is sharp (that is, there exists an intersecting family of size 2^{n-1}).*

Proof. In total, X has 2^n subsets, which can be arranged in pairs of sets that are complements of each other. We have 2^{n-1} such pairs. Out of each pair, $\mathcal{F}$ may contain at most one.

To see that the bound is sharp, take an arbitrary element of X, and consider all subsets of X containing this element. □

Theorem 50 (Erdős-Ko-Rado)**.** *Let $k \leq \frac{n}{2}$, X be a set with $|X| = n$ and $\mathcal{F}$ an intersecting family of k-element subsets of X. Then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1}.$$

Proof. The proof can be found in [Juk], pages 100-101. □

Remark. The theorem above does not hold in the case when $k > n/2$. Indeed, consider the family of all k-element subsets of an n-element set X. This will be intersecting by the pigeonhole principle. The cardinality of this family is $\binom{n}{k}$, which is larger than $\binom{n-1}{k-1}$.

Corollary 51 (Generalized Erdős-Ko-Rado)**.** *Let $\mathcal{F} = A_1, \ldots, A_m \subseteq \{1, \ldots, n\}$ be an intersecting set family with $|A_i| \leq k$ for every i, and suppose $\mathcal{F}$ is a Sperner family, i.e., $A_i \nsubseteq A_j$ for every $i \neq j$. Then $m \leq \binom{n-1}{k-1}$*

Proof. The A_i correspond to elements of the Boolean lattice (the poset of $2^{[n]}$). We can assume that there are sets in $\mathcal{F}$ strictly below the k'th level. Let l be the lowest level that contains such a set. Levels l and $l+1$ induce a bipartite graph that has a matching that contains every vertex on the l'th level (see also the second proof of Sperner's theorem in Lecture 8). Note that no two elements of $\mathcal{F}$ are connected by an edge (because it is a Sperner family). So we can replace each element of $\mathcal{F}$ on the l'th level with its neighbor in the matching, and the resulting family will still satisfy all our conditions. □

Basics of the Probabilistic method. We denote by Ω a probability space, that is, the set consisting of some elements called *elementary events* equipped with a measure p such that
1. $p(A) \geq 0$ for any event (by *event* we mean any union of some elementary events)
2. $p(\Omega) = 1$.
3. $p(A \cup B) = p(A) \cup p(B)$ for any disjoint events A, B.

For simplicity of exposition we work only with discrete probability here, that is, we assume that Ω is finite.

A *random variable* $X : \Omega \to \mathbb{R}$ is just any measurable function that assigns values to elementary events. Note that the measure p doesn't appear in this definition. However, it appears in the next one. If X takes values $x_1, \ldots, x_k$, then the *expectation* $\mathbb{E}(X)$ of X is defined as $\mathbb{E}(X) = \sum_{i=1}^{k} x_i \Pr(X = x_i)$. Note that $\sum_{i=1}^{k} \Pr(X = x_i) = 1$. Informally, it is a weighted average of X with respect to p.

Definition 52. Two events $A, B \subseteq \Omega$ are said to be *independent* if $p(A \cap B) = p(A)p(B)$. Events $A_1, \ldots, A_k$ are said to be independent if $p(\cap_{i \in I} A_i) = \prod_{i \in I} p(A_i)$ for any set $I \subseteq [k]$.

Discrete random variables $X_1, \ldots, X_k$ are said to be independent if the events $X_i = a_i$ are independent for any choice of the a_i's.

Some useful properties:

a. The probability of a union of events $A_1, ..., A_n$ is at most the sum of the probabilities of the events

$$P(A_1 \cup ... \cup A_n) \leq P(A_1) + ... + P(A_n).$$

b. If $A_1, ..., A_n$ are independent discrete random variables, then

$$P(A_1 \cap ... \cap A_n) = P(A_1) \cdot ... \cdot P(A_n).$$

c. The linearity of expectation: If $X_1, ..., X_n$ are (any, not necessary independent) random variables and $a_1, ..., a_n$ are real numbers, then

$$\mathbb{E}(a_1 X_1 + ... + a_n X_n) = a_1 \mathbb{E}[X_1] + ... + a_n \mathbb{E}[X_n].$$

d. If $\mathbb{E}X = k$, then there is at least one elementary event A_1 such that $X(A_1) \geq m$, and, analogously, there is at least one elementary event A_2 such that $X(A_2) \leq m$.

A general framework for the probabilistic method is the following: we are given a finite set of objects Ω and $X : \Omega \to \mathbb{R}$ is a function assigning to each object $A \in \Omega$ a real number. The goal is to show that there is at least one element $A \in \Omega$ for which $X(A)$ is at least a given value m. For this, we define a probability distribution $P : \Omega \to [0, 1]$ and consider the resulting probability space, where X becomes a random variable. Showing that the expected value of X is at least m is enough, since, if this holds, then there exists at least one value $A \in \Omega$ for which $X(A) \geq m$.

A quick application:

Theorem 53. *Every graph $G = (V, E)$ contains a bipartite subgraph with at least $|E|/2$ edges.*

Proof. The proof can be found in [Mat], page 307. □

Also refer to the following:
[**Mat**] 10. Probability and probabilistic proofs.
[**Juk**] 3. Probabilistic Counting.

Lecture 11. LYM inequality in probabilistic terms.

Conditional probability $\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}$.

Proof of LYM inequality using the probabilistic method.

Let $\mathcal{F} \subseteq 2^X$ be a family of subsets of X, where $|X| = n$, such that $A \nsubseteq B, \forall A, B \in \mathcal{F}, A \neq B$. Let $m_k = |\{A \in \mathcal{F} : |A| = k\}|$, that is the number of sets in $\mathcal{F}$ containing k elements.(One can see that $m_1 + ... + m_n = m$).We want to prove that

$$\sum_{i=0}^{n} \frac{m_i}{\binom{n}{i}} \leq 1.$$

Recall that $(2^X, \subseteq)$ is a poset.

First, let us observe that the number of maximal chains is $n!$. This is because a maximum length chain has in it exactly one k-element set for each $0 \leq k \leq n$. Assume you start from the empty set. There are n ways to continue the chain to a 1-element set (think of the number of ways in which we can add one element). For each one-element set, there are $n-1$ ways to continue the chain to a 2-element set, and so on.

We choose randomly with equal probability $1/n!$ one such maximum chain and we define a random variable Y as follows

$$Y = \sum_{F \in \mathcal{F}} Y_F, \text{ where } Y_F = \begin{cases} 1 \text{ if } F \text{ is in the chain} \\ 0 \text{ otherwise} \end{cases}.$$

We know that $\mathbb{E}(Y_F) = P(Y_F = 1) = \frac{|F|!(n-|F|)!}{n!} = 1/\binom{n}{|F|}$. In order to see this, we just count the number of ways can we prolong a chain passing through a set F above and below. Below, there are $|F|!$ possibilities, while above there are $(n-|F|)!$ possibilities. On the other hand, the total number of chains is $n!$, which gives the statement above.

By the linearity of expectation, we obtain

$$\mathbb{E}(Y) = \sum_{F \in \mathcal{F}} \mathbb{E}(Y_F) = \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}}.$$

On the other hand, a chain can contain at most one element of $\mathcal{F}$, so $\mathbb{E}(Y) \leq 1$. Recalling that m_i is the number of i-element subsets in $\mathcal{F}$, we obtain that

$$\mathbb{E}(Y) = \sum_{i=0}^{n} \frac{m_i}{\binom{n}{i}} \leq 1.$$

Definition 54. A collection $\mathcal{F} \subseteq \binom{[n]}{k}$ of k-element subsets of $[n]$ is called a *k-uniform hypergraph.*

Definition 55. A hypergraph $\mathcal{F}$ is *2-colorable* if we can color the elements of $[n]$ by 2 colors such that no member of $\mathcal{F}$ is monochromatic.

Example 56. For $n = 2k-1$, $\mathcal{F} = \binom{[n]}{k}$ is not 2-colorable. Indeed, in any 2-coloring, at least $\lceil n/2 \rceil = k$ elements get the same color, so there is a k-set in $\mathcal{F}$ that is monochromatic.

Here $|\mathcal{F}| = \binom{2k-1}{k} < 2^{2k-1} < 4^k$.

Theorem 57. *Let $\mathcal{F}$ be a k-uniform hypergraph, $|\mathcal{F}| < 2^{k-1}$. Then $\mathcal{F}$ is 2-colorable.*

Proof. Color the elements of $[n]$ independently by green or purple with probability $1/2$. Fix $F \in \mathcal{F}$.

$$\Pr[F \text{ is monochromatic (i.e., "bad")}] = 2 \cdot \frac{1}{2^k} = \frac{1}{2^{k-1}}$$

because it is green with probability $1/2^k$ and purple with probability $1/2^k$.

$$\Pr[\mathcal{F} \text{ has a monochromatic hyperedge } F] \leq \sum_{F \in \mathcal{F}} \Pr[F \text{ is monochromatic}] = \frac{|\mathcal{F}|}{2^{k-1}} < 1$$

So $\Pr[\mathcal{F} \text{ has no monochromatic hyperedge}] > 0$, which implies, in particular, that there is some coloring where no edge is monochromatic. □

This only shows the existence of a coloring. But it can also be turned into an algorithm (Erdős-Selfridge):

Two players, purple and green play and follow the following strategy (green starts):

- for purple (green), a set $F \in \mathcal{F}$ has weight 0 if F already has an element colored purple (green), and it has weight 2^s otherwise, where s is the number of green (purple) elements in F
- $val(x) = \sum_{F \ni x} w(F)$ is the value of an element x
- purple (green) picks a vertex of largest value and colors it purple (green)

It is enough to prove the following claim:

Claim 58. *The total weight of the members of $\mathcal{F}$ for purple (green) is always less than 2^k.*

Indeed, if the game ever produced an edge $F \in \mathcal{F}$ such that all elements of F got colored with the same color, say purple, then this edge would have weight 2^k for green. According to the claim, this never happens, so the algorithm will result in a proper 2-coloring of $\mathcal{F}$.

Proof. We prove the claim for purple. Before purple's first turn, every edge has weight 1 or 2 for purple (depending on whether or not green colored any element of it), so the total weight is less than 2^k. Purple colors an element x such that $val(x)$ is maximum for them. This decreases the total weight for purple by $val(x)$. Then green colors some element y according to their strategy. This increases the total weight for purple by $val(y)$.

But $val(x) \geq val(y)$ by definition, so before purple's next turn, the weight for purple is less than $2^k - val(x) + val(y) \leq 2^k$, which is what we wanted to show. □

Also refer to the following:
[**Mat**] 10. Probability and probabilistic proofs.
[**Juk**] 3. Probabilistic Counting.

Lecture 12. Generating functions.

Generating functions

Example 1. Consider the following combinatorial problem: How many ways are there to pay the amount of 21 francs if we have 6 one-francs coins, 5 two-francs coins, and 4 five-francs coins?

The required number is in fact the number of solutions of the equation

$$x_1 + x_2 + x_3 = 21,$$

with $x_1 \in \{0, 1, 2, 3, 4, 5, 6\}$, $x_2 \in \{0, 2, 4, 6, 8, 10\}$, and $x_3 \in \{0, 5, 10, 15, 20\}$.

In order to compute this, we associate to each variable x_i above a polynomial p_i as follows:

$$p_1(x) = 1+x+x^2+x^3+x^4+x^5+x^6,\ p_2(x) = 1+x^2+x^4+x^6+x^8+x^{10},\ p_3(x) = 1+x^5+x^{10}+x^{15}+x^{20}.$$

The number of solutions of the equation above will be the coefficient of x^{21} in the product $p_1(x) \cdot p_2(x) \cdot p_3(x)$.

Example 2. We prove that $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$. By the binomial theorem, we know that $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$.

Consider the derivative of this polynomial. This gives

$$n(1+x)^{n-1} = \sum_{k=1}^{n} \binom{n}{k} k x^{k-1}.$$

Considering $x = 1$ in the above, we obtain the desired identity.

Example 3. We prove that

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

We know that $(1+x)^n(1+x)^n = (1+x)^{2n}$. Consider the coefficient of x^n in this expression. On the one hand, it is $\binom{2n}{n}$. On the other hand, from the left, we obtain that this coefficient is $\sum_{k=0}^{n} \binom{n}{k}\binom{n}{n-k}$, and since $\binom{n}{k} = \binom{n}{n-k}$, we obtain the desired identity.

Theorem 59. *Let $a_0, a_1, ...$ be a sequence of real numbers. If $|a_k| \leq c^k$ for every k, where c is a positive real constant, then the series*

$$a_0 + a_1 x + a_2 x^2 + ...$$

is convergent for all x with $|x| < 1/c$.

Definition 60. Let $(a_0, a_1, ...)$ be a sequence of real numbers. Then, its *generating function* $a(x)$ is

$$a(x) = a_0 + a_1 x + a_2 x^2 + ...$$

Example 1. The generating function of the sequence $(0, 1, 1/2, 1/3, ...)$ is

$$a(x) = 0 + x + \frac{1}{2}x + \frac{1}{3}x^2 + ... = -\ln(1-x).$$

Example 2. The generating function of the sequence $(1, \frac{1}{1!}, \frac{1}{2!}, \frac{1}{3!}, ...)$ is

$$a(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + ... = e^x.$$

We can sum, multiply generating functions. We can take derivatives and integrals.

Theorem 61 (Generalized binomial theorem)**.** *For every $r \in \mathbb{R}$ and every integer $k \geq 0$, let*

$$\binom{r}{k} = \frac{r(r-1)...(r-k+1)}{k!}.$$

Then, the following holds:

$$(1+x)^r = \binom{r}{0} + \binom{r}{1}x + \binom{r}{2}x^2 + ...$$

for every x with $|x| < 1$.

Refer as well to the following:
[**Mat**] 12. Generating functions.

Lecture 13. Generating functions - The linear algebra method

Examples on generating functions Example 1. $a_{i-1} = i$. Then the generating function $A(x) = 1 + 2x + 3x^2 + \ldots = (1 + x + x^2 + \ldots)' = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}$.

Example 2. $b_{i-1} = i^2$. Arguing in a similar way, one gets that $B(x) = A'(x) - A(x)$.

Example 3. Combinatorial application. Given 10 red balls, 20 blue balls and 30 green balls, in how many ways we can choose 40 balls out of it? It is a coefficient in front of x^{40} in the polynomial

$$p(x) = (1 + x + x^2 + \ldots + x^{10})(1 + x + x^2 + \ldots + x^{20})(1 + x + x^2 + \ldots + x^{30})$$

. Using geometric series formula, we rewrite it in the form

$$p(x) = \frac{1-x^{11}}{1-x}\frac{1-x^{21}}{1-x}\frac{1-x^{31}}{1-x} = \frac{1}{(1-x)^3}(1 - x^{11} - x^{21} - x^{31} + x^{42} + \ldots)$$

and, using the generalized binomial formula for $(1-x)^{-3}$, find the coefficient in front of x^{40}.

Fibonacci sequence

The Fibonacci sequence $(F_n)_{n\geq 0}$ is defined by the following recursive formula:

$$F_0 = 0,\ F_1 = 1,\ F_n = F_{n-1} + F_n, \forall n \geq 2.$$

The sum of the first n numbers of the Fibonacci sequence, is

$$\sum_{k=0}^{n} F_k = F_{n+2} - 1.$$

We want to find an explicit formula for the value of the n-th Fibonacci number. We will present several possible ways to do that.

Method 1.

We will use the generating functions.

Let $F(x)$ denote the generating function of the Fibonacci sequence $(F_0, F_1, ...)$, that is

$$F(x) = F_0 + F_1 x + F_2 x^2 + F_3 x^3 + ...$$

Multiplying $F(x)$ by x, respectively x^2, we obtain that

$$xF(x) = F_0 x + F_1 x^2 + F_2 x^3 + F_3 x^4 + ...$$

$$x^2 F(x) = F_0 x^2 + F_1 x^3 + F_2 x^4 + F_3 x^5 + ...$$

Recall that for every $n \geq 2$, we have $F_n = F_{n-1} + F_{n-2}$ and consider $F(x) - xF(x) - x^2F(x)$. Grouping together the coefficients of x^k for every k, one obtains that

$$F(x)-xF(x)-x^2F(x) = F_0+x(F_1-F_0)+x^2(F_2-F_1-F_0)+x^3(F_3-F_2-F_1)+...+x^k(F_k-F_{k-1}-F_{k-2})+...$$

This implies $F(x) - xF(x) - x^2F(x) = x$ and thus

$$F(x) = \frac{x}{1-x-x^2}.$$

This means, the general term is

$$F_n = \frac{F^{(n)}(0)}{n!},$$

where $F^{(n)}(0)$ is the value in 0 of the n-th derivative of $F(X)$.

We factor $1 - x - x^2$ as $-(x - x_1)(x - x_2)$, where $x_{1,2} = \frac{-1\pm\sqrt{1+4}}{2} = \frac{-1\pm\sqrt{5}}{2}$.

This means

$$F(X) = \frac{x}{1-x-x^2} = \frac{A}{x-x_1} + \frac{B}{x-x_2} = \frac{A(x-x_2)+B(x-x_1)}{-(1-x-x^2)}.$$

From this we obtain that

$$A + B = -1 \text{ and } Ax_2 + Bx_1 = 0.$$

This is a system of two equations with A and B as unknowns, so we can obtain exact values for A and B. One can obtain that:

$$\frac{A}{x - x_1} = \frac{-A/x_1}{1 - x/x_1} = -\frac{A}{x_1} \sum_{n=0} \left(\frac{x}{x_1}\right)^n = -A \sum_{n=0}^{\infty} \frac{x^n}{x_1^{n+1}}$$

$$\frac{B}{x - x_2} = \frac{-B/x_2}{1 - x/x_2} = -\frac{B}{x_2} \sum_{n=0} \left(\frac{x}{x_2}\right)^n = -B \sum_{n=0}^{\infty} \frac{x^n}{x_2^{n+1}}.$$

This implies that the general term F_n is

$$F_n = -\frac{A}{x_1^{n+1}} - \frac{B}{x_2^{n+1}}.$$

Using $F_0 = 0, F_1 = 1$, we obtain the values of A and B, we obtain

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^2 - \left(\frac{1-\sqrt{5}}{2}\right)^2 \right)^2.$$

Another way to interpret the Fibonacci sequence is the following: let S_n denote the number of ways in which one can climb n stairs if allowed to jump one or two stairs at a time. This is the same as to count the number of the solutions of the equation $x_1 + \ldots + x_k = n$ where $x_i \in \{1, 2\}$ and the number k is not fixed. Thus, S_n is the coefficient in front of x^n in the infinite sum $S(x) = 1 + (x + x^2) + (x + x^2)(x + x^2) + \ldots$. So $S(x)$ is the generating function for S_n and, since $S(x)$ is just a geometric progression, we get that $S(x) = \frac{1}{1-x-x^2}$. Next we note that $S_n = F_{n+1}$ and therefore $F(x) = xS(x) = \frac{x}{1-x-x^2}$.

Method 2.

We look first for a geometric series that satisfies $F_n = F_{n-1} + F_{n-2}$, that is $F_n = c \cdot \alpha^n, \forall n$.

This implies that $c\alpha^n = c\alpha^{n-1} + c\alpha^{n-2}$ and thus $\alpha^2 - \alpha - 1 = 0$. Solving this quadratic equation, we get $\alpha_{1,2} = \frac{1 \pm \sqrt{5}}{2}$.

Therefore,

$$F_n = c_1\alpha_1^n + c_2\alpha_2^n = c_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + c_2 \left(\frac{1-\sqrt{5}}{2}\right)^n$$

also satisfies $F_n = F_{n-1} + F_{n-2}$ for any $c_1, c_2 \in \mathbb{R}$.

On the other hand, we can obtain the values of c_1 and c_2 from the following

$$0 = F_0 = c_1 + c_2, \qquad 1 = F_1 = c_1 \frac{1+\sqrt{5}}{2} + c_2 \frac{1-\sqrt{5}}{2}.$$

Substituting the values of c_1 and c_2 in

$$F_n = c_1 \left(\frac{1+\sqrt{5}}{2} \right)^n + c_2 \left(\frac{1-\sqrt{5}}{2} \right)^n ,$$

we obtain the formula for the general term of the Fibonacci sequence.

In general, to solve linear recurrence relations of the form

$$a_{n+k} = c_{k-1} a_{n+k-1} + \ldots + c_0 a_n,$$

we have the following recipe. Denote by $\lambda_1, \ldots, \lambda_s$ the (possibly complex) roots of the equation

$$\lambda^k = c_{k-1} \lambda^{k-1} + \ldots + c_o,$$

where λ_i has multiplicity k_i and $\sum_{i=1}^{s} k_i = k$.

Theorem 62. *A formula for a_n is the solutions to the recurrence above if and only if it has the form $a_n = \sum_{i=1}^{s} P_i(n) \lambda_i^n$, where each $P_i(n)$ is a polynomial of degree $k_i - 1$ with coefficients chosen arbitrarily.*

In particular, if all the multiplicities of the roots are equal to 1, then $s = k$ and the general solution has the form $A_1 \lambda_1^n + \ldots + A_k \lambda_k^n$.

Moreover, for any set of initial values $a_0, \ldots, a_{k-1}$ one can find coefficients of the polynomials $P_i(n)$ in the former case or coefficients A_i in the latter case so that the solution fits to the initial values. Note that in both the case of $P_i(n)$ and A_i the number of coefficients to be determined is equal to k, the number of initial values.

Refer as well to the following:
[**Mat**] 12. Generating functions.
[**Lov**] 4. Fibonacci Numbers.

The oddtown theorem

Assume that in a town there are n inhabitants. We want to create m clubs such that every club has an odd number of members, and any two clubs contain an even number of common members. Our goal is to maximize the value of m. We will prove that $m \leq n$.

Formally, we can state this as the following theorem:

Theorem 63. *Let $A_1, \cdots, A_m$ be distinct subsets of $\{1, \cdots, n\}$, so that the cardinality of each A_i is odd, and the cardinality of each $A_i \cap A_j$ is even. Then $m \leq n$.*

We assign to each set A_i its characteristic vector v_i, defined as

$$v_{ij} = \begin{cases} 1 \text{ if } j \in A_i \\ 0 \text{ if } j \notin A_i \end{cases} .$$

We consider these vectors in $\mathbb{F}_2^n$, where $\mathbb{F}_2$ is the finite field of the residues modulo 2. The goal is to prove that the vectors v_i, with $1 \leq i \leq m$ are linearly independent over the field $\mathbb{F}_2$. This will be enough, since there are at most n linearly independent vectors

over an n-dimensional vector space, and since there are m such characteristic vectors, we obtain that $m \leq n$.

We prove that the vectors v_i are linearly independent, by contradiction. We suppose the contrary, so there is a linear combination $\sum_{i=1}^{m} \alpha_i v_i = 0$ with not all the coefficients α_i equalling zero. Let us observe that, for every i, the value $\|v_i\|^2 = \langle v_i, v_i \rangle$ is the cardinality of the set A_i modulo two, and the value $\langle v_i, v_j \rangle$ is the cardinality of the intersection $A_i \cap A_j$ modulo two for all $i \neq j$. Recall that we are working over the field $\mathbb{F}_2$, so since the cardinality of each set A_i is odd, and the cardinality of every intersection $A_i \cap A_j$ is even, we obtain that $\|v_i\|^2 = 1$ and $\langle v_i, v_j \rangle = 0$. This means that we have

$$\alpha_i = \sum_{j=1}^{m} \alpha_j \langle v_j, v_i \rangle = \left\langle \sum_{j=1}^{m} \alpha_j v_j, v_i \right\rangle = 0 \mod 2,$$

for all $1 \leq i \leq m$. This contradicts the assumption that not all α_i are equal to zero, and thus the inequality $m \leq n$ follows.

Lecture 14. The linear algebra method

Fisher's inequality

Theorem 64 (Fisher's inequality)**.** *Let k be a fixed integers and $A_1, \ldots A_m \subseteq \{1, \ldots n\}$ such that $|A_i \cap A_j| = k$, for all $i, j \in \{1, \ldots, m\}$ with $i \neq j$. Then $m \leq n$.*

Proof of Fisher's inequality.

We prove this theorem using again the linear algebra method.

As before, we prove that the characteristic vectors of the sets A_i, for $1 \leq i \leq m$ are linearly independent over $\mathbb{R}$, the field of real numbers. We denote by v_i the characteristic vector of A_i, and we consider it in $\mathbb{R}^n$. For the sake of contradiction, we assume that these vectors are not linearly independent, so there exist coefficients $\alpha_1, \ldots \alpha_m$, not all zero such that $\sum_{i=1}^{m} \alpha_i v_i = 0$. Since the vector $\sum_{i=1}^{m} \alpha_i v_i$ is zero, its norm is zero, so we obtain that

$$0 = \|\sum_{i=1}^{m} \alpha_i v_i\|^2 = \left\langle \sum_{i=1}^{m} \alpha_i v_i, \sum_{i=1}^{m} \alpha_i v_i \right\rangle = \sum_{i=1}^{m} \alpha_i^2 \|v_i\|^2 + \sum_{1 \leq i \neq j \leq m} \alpha_i \alpha_j \langle v_i, v_j \rangle.$$

The value $\|v_i\|^2$ equals the cardinality of A_i, for each i, and $\langle v_i, v_j \rangle$ equals the cardinality of the intersection $A_i \cap A_j$ (so, it is k). Substituting in the relation above, we obtain that

$$0 = \sum_{i=1}^{m} \alpha_i^2 \|v_i\|^2 + \sum_{i \neq j} k \alpha_i \alpha_j = \sum_{i=1}^{m} \alpha_i^2 (|A_i| - k) + k \sum_{1 \leq i,j \leq m} \alpha_i \alpha_j.$$

Now let us observe that $\sum_{1 \leq i,j \leq m} \alpha_i \alpha_j = \left(\sum_{1 \leq i \leq m} \alpha_i\right)^2$, so

$$0 = \sum_{i=1}^{m} \alpha_i^2 (|A_i| - k) + k \left(\sum_{1 \leq i,j \leq m} \alpha_i \right)^2.$$

Clearly, $|A_i| \geq k$ for all i and $|A_i| = k$ for at most one i, since otherwise the intersection condition would not be satisfied. We distinguish two cases: if none of the sets A_i has cardinality k, then the first sum is zero if and only if all the α_i's are zero, which is a contradiction. Otherwise, one of the sets has cardinality k (we can assume it is A_1), so all the coefficients except α_1 are zero. If α_1 is also zero, then we are done, so we can assume $\alpha_1 \neq 0$. But then the right-hand is greater than 0. This is because the last sum vanishes only if at least two of the coefficients α_i are nonzero, which is again a contradiction.

Therefore, the vectors are linearly independent and thus $m \leq n$, which completes the proof.

Two-distance set

Theorem 65 (Larman-Rogers-Shepard)**.** *Every two-distances set in* $\mathbb{R}^n$ *has at most* $\binom{n}{2} + 3n + 2$ *points.*

Proof.

Let $a_1, \cdots, a_m$ be a two-distance set of distinct points in $\mathbb{R}^n$. Let the distance between two points of our set of points $a_1, \cdots, a_m$ can take only one of two values d_1 or d_2, none of which is zero. We can associate with each point a_i the following polynomial in n real variables $x \in \mathbb{R}^n$:

$$f_i(x) = (|x - a_i|^2 - d_1^2)(|x - a_i|^2 - d_2^2).$$

Then $f_i(a_i) = (d_1 d_2)^2 \neq 0$, but $f_i(a_j) = 0$ for every $j \neq i$. Therefore, these polynomials are linearly independent. It is easy to see that every such polynomial is an appropriate linear combination of the following polynomials

$$(\sum_{i=1}^{n} x_i^2)^2, (\sum_{i=1}^{n} x_i^2)x_j, x_i x_j, x_i, 1 \quad \text{for} \quad i, j = 1, \cdots, n;$$

their number is $1 + n + \left(\binom{n}{2} + n\right) + n + 1 = \binom{n}{2} + 3n + 2$. Thus, the polynomials $f_1, \cdots, f_m$ belong to a linear space of dimension at most $\binom{n}{2} + 3n + 2$. As they are linearly independent, their number m cannot exceed the dimension, completing the proof of the theorem.

Refer as well to [**Mat**]: 13. Applications of linear algebra, and [**Juk**] pages 101-102 and Part III - Linear algebra method.

References

[Lov] - Discrete Mathematics (L. Lovasz, J. Pelikan , K. Vesztergombi);

[Bol] - Combinatorics: Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability (B. Bollobas);

[Mat] - Invitation to Discrete Mathematics, (J. Matousek, J. Nesetril).

[Juk] - Extremal combinatorics, (S. Jukna).