UChicago REU 2013 Apprentice Program **Spring 2013**

Lecture 19: July 26th, 2013

Prof. Babai Scribe: David Kim

1 Oddtown

Recall we have n residents, and we are maximizing m, the number of clubs, under different club rules. Let the club system be $\mathcal{C} = \{C_1, \ldots, C_m\}$, each $C_i \subseteq [n]$.

Rules of Oddtown:

1. $(\forall i)(|C_i| = \text{odd})$.
2. $(\forall i \neq j)(|C_i \cap C_j| = \text{even })$.

Theorem 1.1 (Oddtown Theorem - Elwyn Berlekamp) *If there are n residents and m clubs in Oddtown, then $m \leq n$.*

Proof: Let M be the incidence matrix formed with the incidence vectors of $\mathcal{C}$. If we could show claim 1.2, then the rows of M would be linearly independent, and $m \leq n$.

Claim 1.2 *M has full row rank, i.e.* $\text{rk}(M) = m$.

If $\mathbb{1}_S, \mathbb{1}_T$ are incidence vectors for two clubs $S, T \subseteq [n]$, then $\mathbb{1}_S \cdot \mathbb{1}_T = \mathbb{1}_S^T \mathbb{1}_T = |S \cap T|$. So $MM^T = (m_{ij})$ has

$$m_{ij} = \begin{cases} \mathbb{1}_{C_i} \cdot \mathbb{1}_{C_j} = |C_i \cap C_j| & \text{if } i \neq j \\ \mathbb{1}_{C_i} \cdot \mathbb{1}_{C_i} = |C_i| & \text{if } i = j \end{cases}$$

and by Oddtown rules, MM^T has all diagonal entries odd, and all others even:

$$MM^T = \begin{pmatrix} \ddots & & even \\ & odd & \\ even & & \ddots \end{pmatrix} \tag{1}$$

We want to show $\text{rk}(M) = m$ (M has full row rank). It suffices to show that $\text{rk}(MM^T) = m$ (since $\text{rk}(AB) \leq \text{rk}(A)$). (In fact, we know that $\text{rk}(MM^T) = \text{rk}(M)$, but we do not need this fact right now.) Now $\text{rk}(MM^T) = m \Leftrightarrow \det(MM^T) \neq 0.$, i.e., we need to show that

Therefore, Claim 1.2 will follow from this:

Claim 1.3 $\det(MM^T) \neq 0$ *(MM^T is non-singular).*

(In fact, this is equivalent to Claim 1.2.)

The only thing we know about the entries of MM^T is that they are integers, and we know which entries are even and which entries are odd. So we can tell whether $\det(MM^T)$ is even or odd.

Claim 1.4 $\det(MM^T) =$ *odd.*

For $A, B \in M_n(\mathbb{Z})$ we say that $A \equiv B \mod k$ if every entry of $A - B$ is divisible by k.

Exercise 1.5 *If $A, B \in M_n(\mathbb{Z})$ and $A \equiv B \mod k$, then $\det(A) \equiv \det(B) \mod k$.*

For the proof, we only need the basic facts of modular arithmetic: if $a \equiv b \pmod k$ and $c \equiv d \pmod k$ then $a \pm c \equiv b \pm d \pmod k$ and $ac \equiv bd \pmod k$.

The $n!$-term expansion of the determinant is evaluated using these three oprations only (addition, subtraction, multiplication), so congruence is preserved.

Claim 1.6 $MM^T \equiv I \mod 2$

This follows immediately from (1).

So $\det(MM^T) \equiv \det(I) = 1 \mod 2$ and therefore $\det(MM^T) \neq 0$. ■

Remember Elwyn Berlekamp, not just for the Oddtown Theorem. Berlekamp works in "Algebraic Coding Theory," also called the "Theory of Error-Correcting Codes." Linear algebra and polynomials over finite fields are the key tools of this theory.

2 Finite Fields, Isotropic Vectors & Subspaces

Finite fields, also called Galois Fields, are fields with finitely many elements. The number of elements of a field is called its *order*. The order of every finite field is a prime power, and for every prime power q there is a field of order q, denoted by $GF(q)$ or $\mathbb{F}_q$. $GF(q)$ is unique up to isomorphism.

The fields of prime order, $GF(p)$, are simply the integers modulo p.

The field $GF(p^k)$, where $k > 1$ and p prime, can be represented as follows. Take an irreducible polynomial f of degree k over $GF(p)$. Let $GF(q)$ consist of all the p^k polynomials of degree $\leq k-1$ over $GF(p)$ under the usual addition and modulo f multiplication. It turns out that because of the irrducibility of f, the ring constructed in this manner is a field (there are multiplicative inverses).

Surprisingly, the choice of f does not matter.

Exercise 2.1 *Find an irreducible polynomial of degree 2 over $GF(2)$. Describe the multiplication table in $GF(4)$.*

Exercise 2.2 *For what primes p is the polynomial $x^2 + 1$ irreducible over $GF(p)$?*

Assume F is a (possibly finite) field, and let F^n be the space of $n \times 1$ matrices (column vectors) over F. Let $x \cdot y = x^T y = \sum_{i=1}^n x_i y_i$ be the standard dot product.

Definition 2.3 *We say $x \perp U \leq F^n$ if $(\forall u \in U)(x \perp u)$.*

Definition 2.4 *U-perp is the set of all vectors perpendicular to every vector in U, denoted as $U^\perp = \{x \in F^n : x \perp U\}$.*

Exercise 2.5 *$U^\perp \leq F^n$ is always a subspace.*

Exercise 2.6 *If $U \leq F^n$ (a subspace) $\Rightarrow \dim U + \dim U^\perp = n$.*

Definition 2.7 (Isotropic Vector) *A vector $x \in F^n, x \neq 0$ is isotropic if $x \perp x$, i.e., $\sum_{i=1}^n x_i^2 = 0$.*

Exercise 2.8 *Let $\mathbb{F}_p$ be the field of modulo p residue classes (mod p equivalence classes of integers). This is our standard field of order p. For what values of p does there exist $x \neq 0, x \perp x, x \in \mathbb{F}_p^2$?, i.e., $\exists x_1, x_2 \in \mathbb{Z}$, not both divisible by p, s.t. $p \mid x_1^2 + x_2^2$.*

Definition 2.9 (Totally Isotropic Subspace) *$U \leq V$ is a totally isotropic subspace if $U \perp U$, i.e. $U^\perp \supseteq U$.*

What can we say about the dimension of a totally isotropic space? If U is totally isotropic, then by exercise 2.5, $\dim U \leq \frac{n}{2}$.

3 Eventown

Again, we are given n residents and m clubs, represented by $\mathcal{C} = \{C_1, \ldots, C_m\}$, each $C_i \subseteq [n]$. Recall the rules of Eventown:

- (Rule 0) $C_i \neq C_j$
- (Rule 1) $(\forall i)(|C_i| = \text{even})$
- (Rule 2) $(\forall i \neq j)(|C_i \cap C_j| = \text{even})$

(Rules 1,2 can be combined to a single rule: $(\forall i, j)(|C_i \cap C_j| = \text{even})$.

Theorem 3.1 (Eventown Theorem) *In an Eventown of n residents and m clubs, $m \leq 2^{\lfloor n/2 \rfloor}$.*

An example which achieves the upper bound: create $\frac{n}{2}$ couples and require each couple to join clubs together.

Proof: [Eventown Theorem]

Let's assume now that we are working in the vector space $\mathbb{F}_2^n = \{0,1\}^n$, and let v_i be the incidence vector of C_i. Then we have the following rule: $(\forall i,j)(v_i \cdot v_j = 0$ (because in $\mathbb{F}_2$ every even number is 0 and every odd number is 1). In other words, we have $(\forall i,j)(v_i \perp v_j)$. If $S \subseteq \mathbb{F}_2^n$ is any set of clubs, this is equivalent to $S \perp S \Rightarrow S \subseteq S^{\perp}$.

Lemma 3.2 *Let $U = span(S)$. If $x \perp S \Rightarrow x \perp U$.*

Proof. If $x \perp s$ for every $s \in S$, then $x \perp v$ for every $v \in span(S)$ by distributivity. □

By the above lemma, $S \perp S \Rightarrow S \perp U = span(S) \Rightarrow U \perp U$. So we can say that every maximal set of Eventown clubs is a subspace.

For instance, if we have $n = 6$ and two clubs $A = \{1,2,3,4\}$ and $B = \{3,4,5,6\}$, $A + B = \{1,2,5,6\} = A \triangle B$ (symmetric difference, or $v_A + v_B \mod 2$) is also a possible club satisfying Eventown rules.

But what kind of subspace is U? It is a totally isotropic subspace, as $U \perp U$.

Lemma 3.3 *Let $U \subseteq \mathbb{F}_q^n$ be a subspace in a finite field, where $|\mathbb{F}_q^n| = q^n$. Suppose* $\dim(U) = k$. *What is $|U|$?*

Answer. $|U| = q^k$. Take a basis, say $\{b_1, \ldots, b_k\}$ of U. Then each $u \in U$ is expressed uniquely by $u = \alpha_1 b_1 + \cdots + \alpha_k b_k$. $\Rightarrow$ there are q^k choices of the coefficients.

Summing up, by the two lemmas: if S is a set of Eventown clubs, then $U = span(S)$ is a subspace of Eventown clubs, so it is totally isotropic; therefore $\dim U \leq \frac{n}{2}$ and so $\dim U \leq \lfloor \frac{n}{2} \rfloor$. Finally we conclude that $|U| \leq 2^{\lfloor \frac{n}{2} \rfloor}$. ■

Question: Can you find a totally isotropic subspace of dimension $\frac{n}{2}$?

Example 3.4 *Find an isotropic vector in $\mathbb{C}^2$.*
example: $(1, i)$

Example 3.5 *Find an isotropic vector in $\mathbb{F}_5^2$.*
example: $(1,2)$. *Note:* $2^2 \equiv -1 \mod 5$.

Exercise 3.6 *Find a 50-dimensional, totally isotropic subspace in (a) $\mathbb{F}_5^{100}$ (b) $\mathbb{C}^{100}$.*
example: basis consisting of pairs of coordinates of form example 3.4, 3.5

Definition 3.7 ("maximal" and "maximum") *"maximal" sets are those that are impossible to increase by adding any more elements. "maximum" is simply the largest possible number.*

Exercise 3.8 *Given $n = 100$, create a "maximal" club system of just 2 Oddtown clubs. Find all solutions.*

Answer. Partition the residents into two odd sets. If any other odd-club exists, it intersects each by an even number, which makes it even. (contradiction)

We see that generally speaking, maximal $\neq$ maximum. The First Miracle of Linear Algebra is a 'miracle' in the sense that for linear independence, any maximal set is exactly that which achieves the maximum.

Exercise 3.9 *Every maximal Eventown club system is maximum.*

Ernst Witt: German mathematician (1911-1991); he studied quadratic forms over finite fields and the related maximal totally isotropic subspaces. He was also an active member of the Nazi party, along with Teichmuller, another renowned German mathematician. Ironiously, Witt's thesis advisor was Emmy Noether, the renowned German-Jewish mathematician. He graduated in 1934. From 1933 on, gangs of Nazi students harassed and intimidated Jewish professors, including Noether; they invaded their classes, protesting "Jewish mathematics," making instruction impossible. It seems mathematical brilliance does not preclude prejudice and bigotry. There were also notable examples from the Soviet mathematical hierarchy of the 60s - 80s.

4 Markov Chains

Let's revisit Markov Chains. We are given a set of *states* and an initial probability distribution over the states. (Examples: shuffling a deck of cards (52! states), a frog jumping on a set of water lilies.)

A Markov Chain is a memoryless *stochastic process* where given X_t : state at time t, the probability of reaching a state, say state j, at time $t+1$ is defined by $\mathbb{P}(X_{t+1} = j | X_t = i) = p_{ij}$, dependent only on the current state and the next state in question but not on time or other variables.

The *transition matrix* $T = (p_{ij})$ describes this stochastic process, where every row sums to 1 and each entry $p_{ij} \geq 0$. These matrices are called "stochastic."

Let $q_t \in \mathbb{R}^n$ be the distribution of the frog location at time t. Then the evolution of the Markov Chain is described by:

$$q_{t+1} = q_t T.$$

It follows that $q_t = q_0 T^t$ where q_0 is the initial distribution.

Note that it is impossible to compute this evolution process with computers when the number of states is large, as in card shuffling (T is $52! \times 52!$). Theory is the only possible approach. The basic tools are eigenvalues and eigenvectors.

We always have $T\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, \ldots, 1)^T$ (but a right eigenvector). What is going to be the largest eigenvalue of T?

Exercise 4.1 *If T is a stochastic matrix and $\lambda \in \mathbb{C}$ is an eigenvalue $\Rightarrow |\lambda| \leq 1$.*

A stationary distribution is $q = (q_1, \ldots, q_n)$ s.t. $\forall i, q_i \geq 0$, $\sum_i q_i = 1$, and $qT = q$.

Theorem 4.2 *For all stochastic T, $\exists$ a stationary distribution.*

Theorem 4.3 (The 4th Miracle of Linear Algebra) *Given any matrix, {left eigenvalues} = {right eigenvalues}.*

Proof. $Ax = \lambda x \Rightarrow x^T A^T = \lambda x^T$. Also, $f_A = f_{A^T}$. $\square$

(The 3rd Miracle was that if the columns of a matrix are orthonormal $\Rightarrow$ rows orthonormal).

Now we know that T has a left eigenvector of eigenvalue 1.

Problem: is the left eigenvector nonnegative? This is where the theory of nonnegative matrices, Perron-Frobenius Theory, comes into play.

Theorem 4.4 (Perron–Frobenius) *If $A \geq 0 \Rightarrow \exists$non-negative eigenvector.*

Given $A \geq 0$, we can associate a directed graph $G_A : i \to j$ if $a_{ij} \neq 0$. An important concept in directed graphs is *mutual accessibility* between the vertices of a directed graph.

Definition 4.5 *A directed graph G is strongly connected if every pair of vertices are mutually accessible by directed walks. A strongly connected component is a maximal strongly connected subgraph. A matrix A is irreducible if G_A is strongly connected.*

Exercise 4.6 *If G_A is strongly connected, then the left eigenvector of eigenvalue 1 is unique (up to scaling).*

Exercise 4.7 *If G_A is strongly connected, and if $Ax = \lambda x, x \geq 0 \Rightarrow \forall$ eigenvalues μ, $|\mu| \leq \lambda$.*

The above exercise shows that the greatest positive eigenvalue is also the greatest in absolute value.

Exercise 4.8 *To prove the Perron-Frobenius Theorem, it suffices to prove it for irreducible matrices.*

We show a proof of Perron-Frobenius for $A > 0$ using Brouwer's Fixed Point Theorem:

Let D be a closed disk, and let $f : D \to D$ be a continuous function defined on $D \Rightarrow (\exists x \in D)(f(x) = x)$.

Exercise 4.9 *If the theorem is true for a disk, then it is true for a continuous function defined on a closed square.*

One way to show this: enclose the square with the disk, map each point on the boundary of the disk vertically to the boundary of the rectangle, then to whichever point it gets mapped to by f.

In general, we can extend the theorem to any topologial space which is homeomorphic (mutually continuous bijective mapping) to the closed disk. (Imagine taking a rubber disk, then stretching and compressing it without tearing it.)

Theorem 4.10 *If $A > 0 \Rightarrow \exists$ a positive eigenvector.*

Proof: Let $P = \{v : v \geq 0, ||v|| = 1\}$. Suppose $A > 0$ and define $f : P \to P$ as the mapping $v \mapsto Av \mapsto \frac{Av}{||Av||}$. Note that since $A > 0$, $v \geq 0$, and $v \neq 0$, every entry of Av is positive and $||Av|| \neq 0$. Then by Brouwer's Fixed Point Theorem, $\exists x \in P$ s.t. $f(x) = x$, i.e., $Ax = ||Ax||x$. ■

Importance of "low-dimensional topologies" 3 and 4-dimensional spaces: they behave very differently from higher dimensional spaces. Some problems are easier in higher dimensions.

One of the most famous problems in topology, the Poincaré conjecture, dating to the beginning of the 20th century, concerns the three-dimensional sphere. The analogous problem was solved in higher dimensions (by Steven Smale in 1961 for dimensions ≥ 5 and by Michael Freedman in 1982 for dimension 4). The original Poincaré conjecture was inlcuded among the Clay Institute's seven "Millennium Problems" and became the first one to be solved, by Grigori Perelman (2003), who subsequently declined the Fields Medal as well as the $ 1M reward offered by the Clay Mathematics Institute for solving a Millenium Prize Problem. Other famous Millenium Prize Problems include the Riemann hypothesis (also one of Hilbert's Problems) and the "P vs NP" problem.

5 Comments on applications of the Singular Value Decomposition

Low-rank approximation: key tool in machine learning. Imagine we want to predict the entries of a large matrix such as user preferences (Amazon.com, Netflix, etc.) A low-rank approximation gives a "simple explanation." Principle: "the simplest explanation is the best" ("Occam's razor").

"Principal component analysis:" we project the space on the span of the first few singular vectors and find real meaning to the projections. Key tool in exploratory data analysis.

Projection on the span of just the first two singular vectors: visualization of high-dimensional data.

6 Spectral Theorem revisited: Outline of direct proof.

Let V be a finite dimensional Euclidean space, and $\varphi : V \to V$, symmetric: $(\forall x, y)(\langle \varphi(x), y\rangle = \langle x, \varphi(y)\rangle)$. Need to find orthonormal eigenbasis for φ. Proof steps:

1. Find an eigenvector u
2. Induct on $u^\perp \subseteq V$

Idea: $span(u)$ is a φ-invariant subspace, and therefore $U = u^\perp$ is also an invariant subspace under *symmetric* φ.

How to find u? Take u to be the vector that maximizes the Rayleigh quotient $R_\varphi(x) = \frac{\langle x, \varphi(x)\rangle}{||x||^2}$. Show that this u is an eigenvector using the first exercise of yesterday's lecture: let $x := u + tz$ where $z \perp u$ and set $f(t) = R_\varphi(x)$. You will find that if $z \perp u$ then $z \perp \varphi(u)$ and therefore $\varphi(u) \perp u^\perp$, which implies that $\varphi(u) \in span(u)$, so u is an eigenvector, as desired.