

Linear Algebra in Combinatorics – Bonus

KRIT BOONSIRISETH

April 19, 2022

1 Geometry in $\mathbb{R}^n$

- Let $\vec{v} = [v_1, \dots, v_n]$ and $\vec{w} = [w_1, \dots, w_n]$ be vectors in $\mathbb{R}^n$. Then, the **dot product** $\vec{v} \cdot \vec{w}$ is defined as

$$\vec{v} \cdot \vec{w} = v_1 w_1 + \dots + v_n w_n.$$

- The **norm**, which is a measure of length, is defined as $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$. The alternate notation $|\vec{v}|$ can be used as well, but this notation is normally reserved for $\mathbb{R}$ and $\mathbb{C}$.
- The **distance** between two vectors $\vec{v}, \vec{w}$ is the norm of their difference: $d(\vec{v}, \vec{w}) = \|\vec{v} - \vec{w}\|$.
- The **angle** between two vectors $\vec{v}, \vec{w}$, denoted $\angle(\vec{v}, \vec{w})$, is defined by

$$\cos \angle(\vec{v}, \vec{w}) = \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \|\vec{w}\|}.$$

- We say that two vectors $\vec{v}, \vec{w}$ are **orthogonal**, written as $\vec{v} \perp \vec{w}$, if $\vec{v} \cdot \vec{w} = 0$.
- A basis of $\mathbb{R}^n$ in which any two vectors are orthogonal is called an **orthogonal basis**. If, in addition, each vector has norm equal to 1, the basis is said to be **orthonormal**.

2 Trace

- The **trace** of a square matrix A , denoted $\text{tr } A$, is equal to the sum of elements on the diagonal of A .
- The trace is a linear map, because $\text{tr}(A + B) = \text{tr } A + \text{tr } B$, and $\text{tr}(cA) = c \text{tr } A$ for any scalar c .
- Since the trace only depends on diagonal entries, $\text{tr } A = \text{tr } A^T$.
- The trace of a matrix A is invariant under **changes of basis**, i.e. a way to write A in a different basis. Formally, for a basis matrix B (a matrix whose columns form a basis of V), $\text{tr } A = \text{tr } B^{-1}AB$. This means that the **trace is an inherent property of the linear transformation given by A** , and does not depend on how exactly we express this linear transformation as a matrix. (It may sound confusing now, but it turns out that in many cases we will not want to specify a choice of basis!)
- For $m \times n$ and $n \times m$ matrices A and B , $\text{tr}(AB) = \text{tr}(BA)$. This *does not* imply that $\text{tr}(ABC) = \text{tr}(ACB)$!

3 Determinant

- The **determinant** of a square matrix A , denoted $\det A$, is a scalar value computed from the matrix.
- One way to evaluate a determinant (useful in ‘abstract’ problems, not generally useful for numerically computing determinants) is to take a sum over all permutations σ of $\{1, 2, \dots, n\}$:

$$\det A = \sum_{\sigma \in S_n} (-1)^{\text{sgn } \sigma} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

where the **sign** of a permutation is equal to the number of inversions in it (i.e. the number of $i < j$ such that $\sigma(i) > \sigma(j)$).

For example, the determinant of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

can be computed to be

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

by the following table.

Permutation	Inversions	Sign	Traverse
$1 \mapsto 1, 2 \mapsto 2, 3 \mapsto 3$	0	+	$a_{11}a_{22}a_{33}$
$1 \mapsto 1, 2 \mapsto 3, 3 \mapsto 2$	1	−	$a_{11}a_{23}a_{32}$
$1 \mapsto 2, 2 \mapsto 1, 3 \mapsto 3$	1	−	$a_{12}a_{21}a_{33}$
$1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1$	2	+	$a_{12}a_{23}a_{31}$
$1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 2$	2	+	$a_{13}a_{21}a_{32}$
$1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 1$	3	−	$a_{13}a_{22}a_{31}$

- The determinant is invariant under transposition, that is, $\det A = \det A^T$ for any square matrix A .
- The determinant is multiplicative, that is, $\det AB = \det A \cdot \det B$. This implies that $\det A^{-1} = (\det A)^{-1}$, and also that $\det B^{-1}AB = \det A$, so the determinant is also invariant under changes of basis (and thus a property of *linear transformations*, not just matrices).
- The determinant of a matrix A is nonzero if and only if it is invertible (or equivalently, has full rank).
- Geometrically, the determinant of a matrix A (with real entries) is precisely how much the transformation given by A scales the **signed volume** in $\mathbb{R}^n$.

If $P(\vec{v}_1, \dots, \vec{v}_n)$ is the parallelotope in $\mathbb{R}^n$ spanned by the vectors $\vec{v}_1, \dots, \vec{v}_n$, then the volume of $P(\vec{v}_1, \dots, \vec{v}_n)$ will be equal to $|\det[\vec{v}_1, \dots, \vec{v}_n]|$. This is because this parallelotope is precisely the image of the hypercube $[0, 1]^n$ under the matrix $[\vec{v}_1, \dots, \vec{v}_n]$.

- $\det[\vec{v}_1, \dots, \vec{v}_n]$ is a **multilinear** and **antisymmetric** map $(\mathbb{R}^n)^n \rightarrow \mathbb{R}$. In more understandable terms, it is linear in each of the arguments $\vec{v}_1, \dots, \vec{v}_n$, and if you switch any two arguments, the determinant will be multiplied by -1 .

4 Eigenstuff

- The **characteristic polynomial** of a square $n \times n$ matrix A (with real or complex entries) is given by the determinant $\det(A - \lambda I)$. This is a polynomial in λ of degree exactly n . As you might guess by now, the characteristic polynomial (and eigenvalues, by extension) also does not depend on a choice of basis, because $\det(B^{-1}AB - \lambda I) = \det(B^{-1}(A - \lambda I)B) = \det(A - \lambda I)$.
- $\lambda \in \mathbb{C}$ is said to be an **eigenvalue** of A if $\det(A - \lambda I) = 0$, and the **multiplicity** of an eigenvalue λ is just its multiplicity as a root of the polynomial $\det(A - \lambda I)$. By the Fundamental Theorem of Algebra, an $n \times n$ matrix A will have exactly n eigenvalues, counting multiplicities. We often denote the eigenvalues of a matrix A by $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.
- For an $n \times n$ matrix A with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$,

$$\operatorname{tr} A = \lambda_1 + \lambda_2 + \dots + \lambda_n \quad \text{and} \quad \det A = \lambda_1 \lambda_2 \dots \lambda_n.$$

- A vector $\vec{v}$ is said to be an **eigenvector** to the eigenvalue λ if $A\vec{v} = \lambda\vec{v}$. Note that the existence of such $\vec{v}$ implies λ is an eigenvalue, because we have $(A - \lambda I)\vec{v} = \vec{0}$, which means $\vec{v} \in \ker(A - \lambda I)$ and thus $A - \lambda I$ cannot have full rank.
- A matrix A is said to be **diagonalizable** if there is an invertible matrix B such that $B^{-1}AB$ is a diagonal matrix (i.e. all of its non-diagonal entries are zero).
- If the eigenvectors of an $n \times n$ matrix A spans the vector space $\mathbb{R}^n$ (or $\mathbb{C}^n$), then the matrix $B^{-1}AB$ is diagonal (and hence A is diagonalizable).

5 Spectral graph theory

- A real, symmetric matrix A is always diagonalizable. Consequently, the adjacency matrix of a simple graph is always diagonalizable.
- A (simple) graph G is said to be **d -regular** if each of its vertices has degree exactly d . If $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of the adjacency matrix of G , then the following holds:

- i. $d = \lambda_1 \geq \dots \geq \lambda_n \geq -d$.
- ii. If G is connected, then $\lambda_2 < d$.
- iii. G is bipartite if and only if $\lambda_n = -d$.

- An (n, d, λ) graph is a d -regular graph on n vertices satisfying $\max\{|\lambda_2|, |\lambda_n|\} \leq \lambda$. Generally, the smaller the parameter λ is, the closer the properties of the graph are to the ‘expected’ value.

For example, if G is an (n, d, λ) -graph, then the **expander mixing lemma** states that for any vertex subsets X, Y of G , the number $e(X, Y)$ of pairs $(x, y) \in X \times Y$ such that (x, y) is an edge of G satisfies

$$\left| e(X, Y) - \frac{d}{n}|X||Y| \right| \leq \lambda \sqrt{|X||Y|}.$$

- An **expander graph** is a graph in which, intuitively, a ‘gas cloud’ starting at some small subset of vertices of the graph will expand to fill the graph and uniformize itself ‘quickly’.

There are several different ways to quantify this intuition. (Here the graph G is always assumed to be an n -vertex d -regular graph, and S is a nonempty vertex subset of size at most $n/2$.)

- The **vertex expansion ratio** of G is

$$c := \min_S \frac{|\partial S|}{|S|},$$

where ∂S denotes the set of vertices not in S that are adjacent to at least one vertex in S .

- The **edge expansion ratio** of G is

$$h := \min_S \frac{e(S, V(G) \setminus S)}{|S|}.$$

- The **spectral gap** of G is the gap between the two largest eigenvalues of G :

$$\kappa := d - \lambda_2.$$

- A graph that is a ‘good’ expander should have $c \approx 1$ (this means c is roughly of constant size) and $h, \kappa \approx d$. Setting $h' := h/d$ and $\kappa' := \kappa/d$, we have the following relations:
 - $c \geq h' \geq c/d$. This is easy—try to prove it yourself!
 - **Cheeger’s inequality:** $\sqrt{2\kappa'} \geq h' \geq \kappa'/2$.

6 The Sensitivity Conjecture

If you’ve made it this far (or in fact, even if you just skipped this entire document up to this point) you already know all the tools used in Hao Huang’s proof of the **Sensitivity Conjecture**. In fact, let’s do this proof step-by-step together. Despite how (relatively) simple the proof looks, this was only proven in 2019 (!).

Theorem. (Huang, 2019) *Let G be the n -dimensional hypercube graph, i.e. the graph with vertex set $\{0, 1\}^n$ where two vertices are adjacent if they differ in exactly one coordinate. Then any induced subgraph of G on $2^{n-1} + 1$ vertices has maximum degree at least $\sqrt{n}$.*

1. Define the symmetric matrices A_n of size $2^n \times 2^n$ by $A_0 = [0]$, and

$$A_n = \begin{bmatrix} A_{n-1} & I_{n-1} \\ I_{n-1} & -A_{n-1} \end{bmatrix}$$

for $n \geq 1$. Show that $A_n^2 = nI_{2^n}$.

2. Let B_n be the $2^n \times 2^{n-1}$ matrix

$$\begin{bmatrix} A_{n-1} + \sqrt{n}I_{2^{n-1}} \\ I_{2^{n-1}} \end{bmatrix}.$$

Show that B_n has rank 2^{n-1} and $A_n B_n = \sqrt{n} B_n$.

(Note that B_n is chosen to be precisely the eigenvectors of A_n to the eigenvalue $\sqrt{n}$; the ‘miracle’ in this entire proof is that $\sqrt{n}$ is an eigenvalue of A_n with multiplicity 2^{n-1} .)

3. Identify the rows of B_n with the vertices $(0, \dots, 0), (0, \dots, 0, 1), \dots, (1, \dots, 1)$ of G (essentially counting up in binary). Let H be an induced subgraph of $2^{n-1} + 1$ vertices of G , and let B^* be the $2^{n-1} - 1$ rows of B_n that corresponds to vertices *not* in H . Show that we can find a vector $\vec{x} \in \mathbb{R}^{2^{n-1}}$ such that $B^* \vec{x} = \vec{0}$.

4. Let $\vec{y} = [y_0, \dots, y_{2^n-1}] := B_n \vec{x}$. Observe that the entries of $\vec{y}$ corresponding to vertices outside H are all zero, and that $A_n \vec{y} = \sqrt{n} \vec{y}$. Deduce that if α is such that $|y_\alpha| = \max\{|y_0|, \dots, |y_{2^n-1}|\}$, then the vertex of H corresponding to α has at least $\sqrt{n}$ neighbors in H .

Note that it has been known for a long time that $\sqrt{n}$ is the optimal bound here, as seen by the following construction. Again, let G be the n -dimensional hypercube graph.

Let $n \geq 2$ and let us fix a partition $\{1, \dots, n\} = S_1 \sqcup \dots \sqcup S_k$ of $\{1, \dots, n\}$ into $k \leq \lceil \sqrt{n} \rceil$ non-empty sets which are each of size at most $\lceil \sqrt{n} \rceil$. Define

$$\mathcal{F}_1 = \{T \subseteq \{1, \dots, n\} \mid T \cap S_i \neq \emptyset \text{ for } i = 1, \dots, k\}$$

to be the family of all subsets of $\{1, \dots, n\}$ that intersect each of the sets $S_1, \dots, S_k$ and let

$$\mathcal{F}_2 = \{T \subseteq \{1, \dots, n\} \mid T \cap S_i = \emptyset \text{ for some } i \in \{1, \dots, k\}\}$$

be its complement, i.e. the family of all subsets of $\{1, \dots, n\}$ that are disjoint from the sets $S_1, \dots, S_k$. Furthermore, define

$$\mathcal{F}_{\text{even}} = \{T \subseteq \{1, \dots, n\} \mid |T| \text{ is even}\} \quad \text{and} \quad \mathcal{F}_{\text{odd}} = \{T \subseteq \{1, \dots, n\} \mid |T| \text{ is odd}\},$$

and finally,

$$\mathcal{F} := (\mathcal{F}_1 \cap \mathcal{F}_{\text{even}}) \cup (\mathcal{F}_2 \cap \mathcal{F}_{\text{odd}}) \quad \text{and} \quad \overline{\mathcal{F}} := (\mathcal{F}_1 \cap \mathcal{F}_{\text{odd}}) \cup (\mathcal{F}_2 \cap \mathcal{F}_{\text{even}}).$$

Note that every subset of $\{1, \dots, n\}$ naturally corresponds to a vertex of G (by taking its indicator vector), and hence any family of subsets $\{1, \dots, n\}$ corresponds to an induced subgraph of G

5. Show that every vertex in the induced subgraph $G[\mathcal{F}]$ of G corresponding to the family $\mathcal{F}$ has degree at most $\lceil \sqrt{n} \rceil$. This result also holds for $G[\overline{\mathcal{F}}]$ by the same argument.
6. Show that $|\mathcal{F}|$ and $|\overline{\mathcal{F}}|$ are odd, and that this implies at least one of $|\mathcal{F}|$ and $|\overline{\mathcal{F}}|$ is at least $2^{n-1} + 1$.
7. Deduce that G has an induced subgraph on $2^{n-1} + 1$ vertices with maximum degree at most $\lceil \sqrt{n} \rceil$.

References

This bonus material is largely based on the following:

- Don Knuth's simplified proof of the Sensitivity Conjecture, 2019, <https://www.cs.stanford.edu/~knuth/papers/huang.pdf>
- Yufei Zhao's class on *Graph Theory and Additive Combinatorics*, MIT 2021, <https://yufeizhao.com/gtac/>
- Lisa Sauermann's class on *Algebraic Methods in Combinatorics*, MIT 2022