Algebraic Methods in Combinatorics

Po-Shen Loh

June 2011

1 Linear Algebra review

1.1 Matrix multiplication, and why we care

Suppose we have a system of equations over some field $\mathbb{F}$, e.g.

$$\begin{aligned} 3x_1 + x_2 - 8x_3 &= 1 \\ 9x_1 - x_2 - x_3 &= 2 \end{aligned}$$

The set of ordered triples (x_1, x_2, x_3) that solve the system is precisely the set of 3-element vectors $\mathbf{x} \in \mathbb{F}^3$ that solve the matrix equation

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \text{ where } \mathbf{A} = \begin{bmatrix} 3 & 1 & -8 \\ 9 & -1 & -1 \end{bmatrix}.$$

Suppose that $\mathbf{A}$ is a square matrix. The equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ always has a solution (all zeros). An interesting question is to study when there are no more solutions.

Definition 1 *A square matrix* $\mathbf{A}$ *is* ***nonsingular*** *if* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *has only one solution: the all-zeros vector.*

Often, nonsingular matrices are also called *invertible*, for the following reason.

Theorem 1 *The following are equivalent:*

(i) *The square matrix* $\mathbf{A}$ *is nonsingular.*

(ii) *There exists another matrix, denoted* $\mathbf{A}^{-1}$ *such that* $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$.

Solution:

- (i) to (ii) row-reduction, find that A always has $RREF = identity$
- then can always solve $Ax = (anything)$
- so can for example solve $Ax = I$, by doing it column by column.
- now we know there is some B such that $AB = I$.
- also, the row operations themselves were some left multiplication of matrices
- so we also have some C so that $CA = I$.
- Then $B = IB = CAB = CI = C$, so they are the same.
- (ii) to (i): just multiply by inverse.

For non-square matrices, the most important fact is the following:

Theorem 2 *If* $\mathbf{A}$ *has more columns than rows, then the equation* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *has more solutions than just the all-zeros vector.*

Solution: RREF, run out of rows before columns, so we don't have the sentinel 1 in every column. This enables us to choose arbitrary nonzero values for all non-sentinel columns, and then still read off a valid solution by backtracking the sentinel columns.

1.2 Vectors

A vector is typically represented as an arrow, and this notion is widely used in Physics, e.g., for force, velocity, acceleration, etc. In Mathematics, vectors are treated more abstractly. In this lecture, we will mostly use concrete vectors, which one can think of as columns of numbers (the coordinates). The fundamental operation in Linear Algebra is the linear combination. It is called "linear" because vectors are not multiplied against each other.

Definition 2 *Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, and real coefficients $c_1, \ldots, c_k$, the sum $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ is called a **linear combination**.*

It's intuitive that two vectors "usually" do not point in the same direction, and that three vectors "usually" do not all lie in the same plane. One can express both of the previous situations as follows:

Definition 3 *Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a collection of vectors. One says that they are **linearly dependent** if:*

- *One of the vectors can be expressed as a linear combination of the others.*
- *There is a solution to $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, using real numbers $c_1, \ldots, c_k$, where not all of the c_i's are zero.*

Solution:

- (i) to (ii): swap signs on the linear combination.
- (ii) to (i): pick a nonzero, then shift all others to other side, and divide by the nonzero coefficient.

A fundamental theorem in Linear Algebra establishes that it is impossible to have "too many" linearly independent vectors. This is at the core of many Combinatorial applications.

Theorem 3 *The maximum possible number of linearly independent vectors in $\mathbb{R}^n$ is n.*

Solution: Suppose we have more than n, say $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$. Try to solve for the coefficients. This sets up a matrix equation with more columns than rows, which by previous we know to have a nontrivial solution.

2 Combinatorics of sets

We begin with a technical lemma.

Lemma 1 *Let $\mathbf{A}$ be a square matrix over $\mathbb{R}$, for which all non-diagonal entries are all equal to some $t \geq 0$, and all diagonal entries are strictly greater than t. Then $\mathbf{A}$ is nonsingular.*

Proof. If $t = 0$, this is trivial. Now suppose $t > 0$. Let $\mathbf{J}$ be the all-ones square matrix, and let $\mathbf{D} = \mathbf{A} - t\mathbf{J}$. Note that $\mathbf{D}$ is nonzero only on the diagonal, and in fact strictly positive there. We would like to solve $(t\mathbf{J} + \mathbf{D})\mathbf{x} = \mathbf{0}$, which is equivalent to $\mathbf{D}\mathbf{x} = -t\mathbf{J}\mathbf{x}$. Let s be the sum of all elements in $\mathbf{x}$, and let the diagonal entries of $\mathbf{D}$ be $d_1, \ldots, d_n$, in order. Then, we have $d_i x_i = -ts \Rightarrow x_i = -(t/d_i)s$. But since t and d_i are both strictly postive, this forces every x_i to have opposite sign from s, which is impossible unless all $x_i = 0$. Therefore, $\mathbf{A}$ is nonsingular.

Solution: ALTERNATE:

Let $\mathbf{J}$ be the all-ones square matrix, and let $\mathbf{D} = \mathbf{A} - t\mathbf{J}$. Note that $\mathbf{D}$ is nonzero only on the diagonal, and in fact strictly positive there, so it is a positive definite matrix. Also, $\mathbf{J}$ is well-known to be positive semidefinite (easy to verify by hand), so $\mathbf{A}$ is positive definite. In particular, this means that $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{0}$ only if $\mathbf{x} = 0$, implying that $\mathbf{A}\mathbf{x} = \mathbf{0}$ only for $\mathbf{x} = \mathbf{0}$. This is equivalent to $\mathbf{A}$ being nonsingular. □

Now try the following problems. The last two come from *102 Combinatorial Problems*, by T. Andreescu and Z. Feng.

1. (A result of Bourbaki on finite geometries; also appeared in St. Petersburg Olympiad.) Let X be a finite set, and let $\mathcal{F}$ be a family of distinct **proper** subsets of X. Suppose that for every pair of distinct elements in X, there is a unique member of $\mathcal{F}$ which contains both elements. Prove that $|\mathcal{F}| \geq |X|$.

   Solution: Let $X = [n]$ and $\mathcal{F} = \{A_1, \ldots, A_m\}$. We need to show that $n \leq m$. Define the $m \times n$ incidence matrix $\mathbf{A}$ over $\mathbb{R}$ by putting 1 in the i-th row and j-th column if $j \in A_i$. Consider the product $\mathbf{A}^T\mathbf{A}$, which is an $n \times n$ matrix. For $i \neq j$, its entry at (i, j) is precisely 1.

   Also, the diagonal entries are strictly larger than 1, because if some element $j \in X$ belongs to only one set $A_k \in \mathcal{F}$, then the condition implies that every element $i \in X$ is also in A_k, contradicting requirement that A_k be **proper**.

   Therefore, $\mathbf{A}^T\mathbf{A}$ is nonsingular by Lemma 1. But if $\mathbf{A}$ has more rows than columns, then it would have some $\mathbf{x} \neq \mathbf{0}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$, hence $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{0}$. Therefore, $\mathbf{A}$ doesn't have more rows than columns, i.e., $n \leq m$.

2. (Fisher's inequality) Let $\mathcal{C} = \{A_1, \ldots, A_r\}$ be a collection of distinct subsets of $\{1, \ldots, n\}$ such that every pairwise intersection $A_i \cap A_j$ $(i \neq j)$ has size t, where t is some fixed integer between 1 and n inclusive. Prove that $|\mathcal{C}| \leq n$.

   Solution: Consider the $n \times r$ matrix $\mathbf{A}$, where the i-th column of $\mathbf{A}$ is the characteristic vector of A_i. Then, $\mathbf{A}^T\mathbf{A}$ is a $r \times r$ matrix, all of whose off-diagonal entries are t. We claim that the diagonal entries are all $> t$. Indeed, if there were some $|A_i|$ which were exactly t, then the structure of $\mathcal{C}$ must look like a "flower," with one set A_j of size t, and all other sets fully containing A_j and disjointly partitioning the elements of $[n] \setminus A_j$ among them. Any such construction has size at most $1 + (n - t) \leq n$, so we would already be done.

   Therefore, $\mathbf{A}^T\mathbf{A}$ is nonsingular by Lemma 1, and the previous argument again gives $r \leq n$.

3. Let $A_1, \ldots, A_r$ be a collection of distinct subsets of $\{1, \ldots, n\}$ such that all $|A_i|$ are even, and also all $|A_i \cap A_j|$ are even for $i \neq j$. How big can r be, in terms of n?

   Solution: Arbitrarily cluster $[n]$ into pairs, possibly with one element left over. Then take all possible subsets where we never separate the pairs; this gives r up to $2^{\lfloor n/2 \rfloor}$.

   But this is also best possible. Suppose that S is the set of characteristic vectors of the sets in the extremal example. The condition translates into S being self-orthogonal. But $S \perp S \Rightarrow \langle S \rangle \perp \langle S \rangle$, so extremality implies that S is in fact an entire linear subspace, which is self-orthogonal (i.e., $S \subset S^\perp$).

   We have the general fact that for any linear subspace, $\dim S^\perp = n - \dim S$. This is because if $d = \dim S$, we can pick a basis $v_1, \ldots, v_d$ of S, and write them as the rows of a matrix $\mathbf{A}$. Then, the kernel of $\mathbf{A}$ is precisely $S^\perp$, but any kernel has dimension equal to n minus the dimension of the row space (d).

   Therefore, $S \subset S^\perp$ implies that $\dim S \leq \dim S^\perp = n - \dim S$, which forces $\dim S \leq \lfloor n/2 \rfloor$, so we are done.

4. What happens in the previous problem if we instead require that all $|A_i|$ are odd? We still maintain that all $|A_i \cap A_j|$ are even for $i \neq j$.

   Solution: Answer: $r \leq n$. Work over $\mathbb{F}_2$. The characteristic vectors v_i of the A_i are orthonormal[1], so they are linearly independent: given any dependence relation of the form $\sum c_i v_i = \mathbf{0}$, we can dot product both sides with v_k and conclude that $c_k = 0$. Thus, there can only be $\leq n$ of them.

   ALTERNATE: Let $\mathbf{A}$ be the $n \times r$ matrix where the columns are the characteristic vectors of the A_i. Then $\mathbf{A}^T\mathbf{A}$ equals the $r \times r$ identity matrix, which is of course of full rank r. Thus $r = \text{rank}(\mathbf{A}^T\mathbf{A}) \leq \text{rank}(\mathbf{A}) \leq n$.

5. Prove that if all codegrees[2] in a graph on n vertices are odd, then n is also odd.

[1] Strictly speaking, this is not true, because there is no positive definite inner product over $\mathbb{F}_2$. However, if one carries out the typical proof that orthonormality implies linear independence, it still works with the mod-2 dot product.

[2] The *codegree* of a pair of vertices is the number of vertices that are adjacent to both of them.

Solution: First we show that all degrees are even. Let v be an arbitrary vertex. All vertices $w \in N(v)$ have odd codegree with v, which means they all have odd degree in the graph induced by $N(v)$. Since the number of odd-degree vertices in any graph must always be even, we immediately find that $|N(v)|$ is even, as desired.

Let A be the adjacency matrix. Then $A^T A = J - I$. But consider right-multiplying by $\mathbf{1}$. $A\mathbf{1} = \mathbf{0} \Rightarrow A^T A\mathbf{1} = \mathbf{0}$ and $I\mathbf{1} = \mathbf{1}$, so we need to have $J\mathbf{1} = \mathbf{1}$, which implies that n is odd.

ALTERNATE ENDING: Now, let $S = \{\mathbf{1}, v_1, \ldots, v_n\}$ be the set of $n+1$ vectors in $\mathbb{F}_2^n$ where $\mathbf{1}$ is the all-ones vector and v_i is the characteristic vector of the neighborhood of the i-th vertex. There must be some nontrivial linear dependence $b\mathbf{1} + \sum_i a_i v_i = 0$. But note that if we take the inner product of this equation with v_k, we obtain $\sum_{i \neq k} a_i = 0$ because $\mathbf{1} \cdot v_k = 0 = v_k \cdot v_k$ and $v_i \cdot v_k = 1$ for $i \neq k$. Hence all the a_i are equal. Yet if they are all zero, then b is also forced to be zero, contradicting the nontriviality of this linear combination. Therefore, all a_i are 1, and the equation $\sum_{i \neq k} a_i = 0$ forces $n - 1$ to be even, and n to be odd.

6. (Introductory Problem 38) There are $2n$ people at a party. Each person has an even number of friends at the party. (Here, friendship is a mutual relationship.) Prove that there are two people who have an even number of common friends at the party.

   Solution: Let A be adjacency matrix. Suppose for contradiction that every pair of people has an odd number of common friends. Then over $\mathbb{F}_2$, we have $A^T A = J - I$, where J is the all-ones matrix and I is the identity. Since all degrees even, $A\mathbf{1} = 0$. Hence $A^T A\mathbf{1} = \mathbf{0}$. But $J\mathbf{1} = \mathbf{0}$ because J is a $2n \times 2n$ matrix, and $I\mathbf{1} = \mathbf{1}$. Thus we have $\mathbf{0} = A^T A\mathbf{1} = (J - I)\mathbf{1} = \mathbf{1}$, contradiction.

7. (Advanced Problem 49) A set T is called *even* if it has an even number of elements. Let n be a positive even integer, and let $S_1, \ldots, S_n$ be even subsets of the set $\{1, \ldots, n\}$. Prove that there exist some $i \neq j$ such that $S_i \cap S_j$ is even.

   Solution: Let A be $n \times n$ matrix over $\mathbb{F}_2$ with columns that are the characteristic vectors of the S_i. Then $A^T A = J - I$, but A is singular because $A^T\mathbf{1} = 0$. Square the equation. We have $(J-I)(J-I) = J^2 - 2J + I$ since I, J commute. But n is even, and we are in $\mathbb{F}_2$, so it is just I, and we get $A^T A A^T A = I$. Contradicts singularity of A.

   (Uses A nonsingular implies A^T nonsingular. Indeed, we need $(AB)^T = B^T A^T$. So, in particular, if A had inverse B, then we have a matrix B^T such that it is left inverse of A^T. In particular, whenever we go to solve $A^T x = 0$, we can left-multiply by B^T, and get $x = B^T 0 = 0$, so no nontrivial solutions.)

   ALTERNATE: Singularity implies $\det A^T A = (\det A)^2 = 0$. However, $\det(J - I)$ is precisely the parity of D_n, the number of derangements of $[n]$. It remains to prove that for even n, D_n is odd. But this follows from the well-known recursion $D_n = (n-1)(D_{n-1} + D_{n-2})$, which can be verified by looking at where the element n is permuted to.

3 Bonus problems (not all linear algebra)

1. (Caratheodory.) A *convex combination* of points x_i is defined as a linear combination of the form $\sum_i \alpha_i x_i$, where the α_i are non-negative coefficients which sum to 1.

   Let X be a finite set of points in $\mathbb{R}^d$, and let $\mathrm{cvx}(X)$ denote the set of points in the convex hull of X, i.e., all points expressible as convex combinations of the $x_i \in X$. Show that each point $x \in \mathrm{cvx}(X)$ can in fact be expressed as a convex combination of only $d+1$ points of X.

   Solution: Given a convex combination with $d+2$ or more nonzero coefficients, find a new vector with which to perturb the nonzero coefficients. Specifically, seek $\sum_i \beta_i x_i = 0$ and $\sum_i \beta_i = 0$, which is $d+1$ equations, but with $d+2$ variables β_i. So there is a non-trivial solution, and we can use it to reduce another α_i coefficient to zero.

2. (Radon.) Let A be a set of at least $d+2$ points in $\mathbb{R}^d$. Show that A can be split into two disjoint sets $A_1 \cup A_2$ such that $\mathrm{cvx}(A_1)$ and $\mathrm{cvx}(A_2)$ intersect.

Solution: For each point, create an $\mathbb{R}^{d+1}$-vector v_i by adding a "1" as the last coordinate. We have a non-trivial dependence because we have at least $d+2$ vectors in $\mathbb{R}^{d+1}$, say $\sum_i \alpha_i v_i = 0$. Split $A = A_1 \cup A_2$ by taking A_1 to be the set of indices i with $\alpha_i \geq 0$, and A_2 to be the rest.

By the last coordinate, we have

$$\sum_{i \in A_1} \alpha_i = \sum_{i \in A_2} (-\alpha_i) .$$

Let Z be that sum. Then if we use α_i / Z as the coefficients, we get a convex combination from A_1 via the first d coordinates, which equals the convex combination from A_2 we get by using $(-\alpha_i)/Z$ as the coefficients.

3. (Helly.) Let $C_1, C_2, \ldots, C_n$ be convex sets of points in $\mathbb{R}^d$, with $n \geq d+1$. Suppose that every $d+1$ of the sets have a non-empty intersection. Show that all n of the sets have a non-empty intersection.

   Solution: Induction on n. Clearly true for $n = d+1$, so now consider $n \geq d+2$, and assume true for $n-1$. Then by induction, we can define points a_i to be in the intersection of all C_j, $j \neq i$. Apply Radon's Lemma to these a_i, to get a split of indices $A \cup B$.

   Crucially, note that for each $i \in A$ and $j \in B$, the point a_i is in C_j. So, each $i \in A$ gives $a_i \in \bigcap_{j \in B} C_j$, and hence the convex hull of points in A is entirely contained in all C_j, $j \in B$.

   Similarly, the convex hull of points in B is entirely contained in all C_j, $j \in A$. Yet Radon's Lemma gave intersecting convex hulls, so there is a point in both hulls, i.e., in all C_j, $j \in A \cup B = [n]$.

4. (From Peter Winkler.) The 60 MOPpers were divided into 8 teams for Team Contest 1. They were then divided into 7 teams for Team Contest 2. Prove that there must be a MOPper for whom the size of her team in Contest 2 was strictly larger than the size of her team in Contest 1.

   Solution: In Contest 1, suppose the team breakdown was $s_1 + \cdots + s_8 = 60$. Then in the i-th team, with s_i people, say that each person did $\frac{1}{s_i}$ of the work. Similarly, in Contest 2, account equally for the work within each team, giving scores of $\frac{1}{s_i'}$.

   However, the total amount of work done by all people in Contest 1 was then exactly 8, and the total amount of work done by all people in Contest 2 was exactly 7. So somebody must have done strictly less work in Contest 2. That person saw

   $$\frac{1}{s_i'} < \frac{1}{s_i} ,$$

   i.e., the size of that person's team on Contest 2 was strictly larger than her team size on Contest 1.

5. (MOP 2007/4/K2.) Let S be a set of 10^6 points in 3-dimensional space. Show that at least 79 distinct distances are formed between pairs of points of S.

   Solution: Zarankiewicz counting for the excluded $K_{3,3}$ in the unit distance graph. This upper-bounds the number of edges in each constant-distance graph, and therefore lower-bounds the number of distinct distances.

6. (MOP 2007/10/K4.) Let S be a set of $2n$ points in space, such that no 4 lie in the same plane. Pick any n^2+1 segments determined by the points. Show that they form at least n (possibly overlapping) triangles.

   Solution: In fact, every $2n$-vertex graph with at least n^2+1 edges already contains at least n triangles. No geometry is needed.

7. (Sperner capacity of cyclic triangle, also Iran 2006.) Let A be a collection of vectors of length n from $\mathbb{Z}_3$ with the property that for any two distinct vectors $a, b \in A$ there is some coordinate i such that $b_i = a_i + 1$, where addition is defined modulo 3. Prove that $|A| \leq 2^n$.

   Solution: For each $a \in A$, define the $\mathbb{Z}_3$-polynomial $f_a(\mathbf{x}) := \prod_{i=1}^n (x_i - a_i - 1)$. Observe that this is multilinear. Clearly, for all $a \neq b \in A$, $f_a(b) = 0$, and $f_a(a) \neq 0$; therefore, the f_a are linearly independent, and bounded in cardinality by the dimension of the space of multilinear polynomials in n variables, which is 2^n.