Problem Set Solutions for

Gems of Linear Algebra

Nandana Madhukara
nandana.madhukara@gmail.com

Contents

1 Week 1

(1) Let $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ be linearly independent vectors over $\mathbb{F}$. Assume $A\mathbf{x}_1, A\mathbf{x}_2, ..., A\mathbf{x}_n$ are linearly dependent vectors over $\mathbb{F}$. Therefore there exists weights $c_1, c_2, ..., c_n$ such that

$$c_1 A\mathbf{x}_1 + c_2 A\mathbf{x}_2 + \cdots + c_n A\mathbf{x}_n = \mathbf{0}.$$

Multiplying by B on the left gives us

$$c_1 BA\mathbf{x}_1 + c_2 BA\mathbf{x}_2 + \cdots c_n BA\mathbf{x}_n = \mathbf{0}$$

which reduces down to

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = \mathbf{0}.$$

This is a contradiction so $A\mathbf{x}_1, A\mathbf{x}_2, ..., A\mathbf{x}_n$ are linearly independent.

Since we have n linearly independent vectors, these vectors span $\mathbb{F}$ so there exist weights $c_1, c_2, ..., c_n$ such that

$$\mathbf{u} = c_1 A\mathbf{x}_1 + c_2 A\mathbf{x}_2 + \cdots + c_n A\mathbf{x}_n$$

for all vectors $\mathbf{u}$ over $\mathbb{F}$. Since A is a linear transformation, we can "factor" it out to get

$$\mathbf{u} = A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n) = A\mathbf{v}$$

where $\mathbf{v}$ is some vector over $\mathbb{F}$. Therefore,

$$AB\mathbf{u} = AB(A\mathbf{v}) = A(BA)\mathbf{v} = A\mathbf{v} = \mathbf{u}$$

for all vectors $\mathbf{u}$. Therefore

$$AB = I_n.$$

(2) Notice that $\mathbf{v}_i\mathbf{w}_i$ is an $m \times n$ matrix where each column is of the form $c\mathbf{v}$ where c is some scalar in $\mathbb{F}$. Therefore each column of $A = \sum_{i=1}^{r} = \mathbf{v}_i\mathbf{w}_i$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$. This means that the column space of A is spanned by $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$ which means that $\text{rank}(A) \leq r$.

(3) If $\text{rank}(A) = n$, all columns of A are linearly independent. Now $A\mathbf{x}$ is just a linear combination of the columns so since the columns are linearly independent, $\mathbf{x} = \mathbf{0}$ is the only solution to $A\mathbf{x} = \mathbf{0}$. This means that dimension of the nullspace is 0 so the rank-nullity theorem holds.

Since $\text{rank}(A) \leq \min(m, n)$, the only case left to consider is $\text{rank}(A) = r < n$. Therefore, there are r linearly independent rows so there are $n - r$ free variables. This means that every vector in the nullspace can be expressed as a linear combination of $n - r$ vectors so the dimension of the nullspace is $n - r$. Therefore, the sum of the rank and the nullity is again n.

(4) We first prove that $\text{nullity}(A) + \text{nullity}(B) \geq \text{nullity}(AB)$. Let the basis of $\ker B$ be $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r\}$. Since $\ker B \subseteq \ker AB$, we can write the basis of $\ker AB$ as

$$\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r, \mathbf{v}_{r+1}, ..., \mathbf{v}_s\}.$$

Now we can see that $\{B\mathbf{v}_{r+1}, ..., B\mathbf{v}_s\}$ are linearly independent.

2 Week 2

(1) Notice that it suffices to prove that each $\lambda_i = 0$ if for all $k \in [n]$,

$$\sum_{i=1}^{n} \lambda_i^k.$$

We can prove this with induction on n. When $n = 1$, it is clear that λ_1 must be 0 so the base case is proven. Now we go on to the inductive step. If any one of the $\lambda_i = 0$ for all k, then we can use the inductive hypothesis on the other λ_j's to finish the proof. Otherwise, consider the Vandermonde Matrix:

$$V = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \cdots & a_n \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \lambda_3^{k-1} & \cdots & \lambda_n^{k-1} \end{pmatrix}.$$

Now our criterion about the sum of the kth powers of the λ_i's always being 0 reduces down to

$$V \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \cdots & a_n \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \lambda_3^{k-1} & \cdots & \lambda_n^{k-1} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{pmatrix} = \mathbf{0}.$$

This means that V so the determinant of V is 0. Since V is a Vandermonde Matrix, by Corollary 2.2, we see that there must exist i, j such that $\lambda_i = \lambda_j$. Now we can use the inductive hypothesis on all the other λ_k's excluding λ_j and we are done.

(2) Let us try to find a path-counting problem such that the number of ways of going from vertex A_i to vertex B_j is $\binom{a_i}{b_j}$. We define our graph with vertices being lattice points and the edges connecting adjacent vertices (excluding diagonal adjacencies). Our paths consists of stepping up or to the right. Next we assign the vertices A_i's and B_j's to the lattice points as the following: A_i is assigned to $(-a_i, 0)$ and B_j is assigned $(-b_j, b_j)$. Notice that the number of ways of getting from A_i to B_j is the number ways to step up or to the right starting from $(-a_i, 0)$ and ending at $(-b_j, b_j)$. The length of this rectangle is $a_i - b_j$ and the height is b_j, so the number of paths is $\binom{a_i - b_j + b_j}{b_j} = \binom{a_i}{b_j}$. By the Lindström–Gessel–Viennot Lemma, the determinant of M is the number of families of non-intersecting paths which must be positive.

(3) We can follow the same approach as the previous problem to get lattice points the A_i's are assigned to are $(i-1, 0)$ and the B_j's are assigned to $(m-j+1, j-1)$. Now we count the number of families non-intersecting paths. We know that A_1 must go to B_1 since if it went to any other B_j, then B_1 would be blocked off. Similarly, A_2 has to go to B_2 or going to another B_j would block off B_j so in general, we see that A_i goes to B_i. However this means that the only family of non-intersecting paths is sideway L's that form shells like a Matryoshka doll. This means that the determinant is just $\boxed{1}$.

3 Week 3

(1) We must prove that $M_{ii} = C_i C_i^T$. This means that the ith row in C_i is missing and the ith column in C_i^T is missing. The row missing corresponds to the ith row missing in M_{ii} and the missing column corresponds to the ith column missing in M_{ii}. This can be seen by the definition of matrix multiplication.

(2) Let us first consider m_{ii} which is just the dot product of the ith row vector of C and the ith column vector of C^T but this is just the same vector. The jth entry of this vector is ± 1 if edge j contains vertex i and 0 if it doesn't. Therefore. dot producting this vector with itself gives us the number of entries that are nonzero which is the number of edges connected to i or the degree of i.

Now consider vertex i and j such that $i \neq j$. Now the m_{ij} is the dot product of the ith row and jth row of C. Now if i and j are not connected, then ± 1's of row i will not line up with the ± 1's of row j so the dot product will be 0. Now if i and j are connected, then a -1 of row i will align with a 1 of row j or vice versa to get a dot product of -1. Therefore m_{ij} is -1 if i and j are connected and 0 if i and j are not connected.

(3) Consider some row i of A. We know that

$$\det(A) = \sum_{j=1}^{n+1} a_{ij}(-1)^{i+j} A_{ij} = 0$$

since the sum of rows and columns is 0. Notice that

$$\sum_{j=1}^{n+1} a_{ij} = 0$$

so

$$(-1)^{i+1} A_{i1} = (-1)^{i+2} A_{i2} = \cdots = (-1)^{i+n+1} A_{i(n+1)}$$

for any row i. Similarly, if we do a cofactor expansion along columns instead of rows, we get

$$(-1)^{1+j} A_{1j} = (-1)^{2+j} A_{2j} = \cdots = (-1)^{n+1+j} A_{(n+1)j}$$

so all cofactors must be equal no matter with i and j are.

(4) If we remove the edge between vertex 1 and vertex n, the Laplacian of this graph would be

$$M = \begin{pmatrix} n-2 & -1 & -1 & \cdots & 0 \\ -1 & n-1 & -1 & \cdots & -1 \\ -1 & -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -1 & -1 & \cdots & n-2 \end{pmatrix}.$$

Removing the first row and column gives us

$$M_{11} = \begin{pmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-2 \end{pmatrix}.$$

Adding rows 2 to $n-1$ to row 1 gives us

$$\begin{pmatrix} 1 & 1 & \cdots & 0 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-2 \end{pmatrix}.$$

Adding the first row to each of the other rows gives us

$$\begin{pmatrix} 1 & 1 & \cdots & 0 \\ 0 & n & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n-2 \end{pmatrix}.$$

This is just an upper triangular matrix so its determinant is $\boxed{n^{n-3}(n-2)}$.

(5) The Laplacian matrix of this graph is

$$M = \begin{pmatrix} n & -1 & -1 & -1 & \cdots & -1 \\ -1 & 3 & -1 & 0 & \cdots & -1 \\ -1 & -1 & 3 & -1 & \cdots & 0 \\ -1 & 0 & -1 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & 0 & 0 & \cdots & 3 \end{pmatrix}.$$

Now we see

$$M_{11} = \begin{pmatrix} 3 & -1 & 0 & \cdots & -1 \\ -1 & 3 & -1 & \cdots & 0 \\ 0 & -1 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 3 \end{pmatrix}$$

and we must take the determinant of this. I have tried doing row operations but I can't seem to find a way to find the determinant.

4 Week 4

(1) We can go through the same process as Theorem 1.1 but we hit a problem when l_{ii} is 0. However, we can just make all the entries of U in row i just 0. Now the main diagonal of U is always 1 so $\det U = 1 \neq 0$ meaning that U is nonsingular. Therefore at least one of L and U is nonsingular. Now we know that the l_{ii}'s are nonzero iff A is invertible. Since A being invertible is equivalent to $k = n$ and the l_{ii}'s being nonzero is equivalent to L being nonsingular, we know that L and U are nonsingular if and only if $k = n$.

(2) We claim that every lower triangular matrix L can be factored as $L'D$ where L is a lower triangular matrix with diagonal entries of 1 and D is a diagonal matrix. Notice that multiplying by a diagonal matrix on the right just scales column j in L' by the nonzero entry of column j in D. Therefore, we can scale each column down by the the entry in that column that is in the diagonal to get L' which must have diagonal entries of 1. Now we make the diagonal entries of D the diagonal entries of L and we are done.

In a similar fashion, we claim that every upper triangular matrix U can be factored as DU' where U is a upper triangular matrix with diagonal entries of 1 and D is a diagonal matrix. Notice that multiplying by a diagonal matrix on the left just scales row i in U' by the nonzero entry of row i in D. Therefore, we can scale each row down by the the entry in that row that is in the diagonal to get U' which must have diagonal entries of 1. Now we make the diagonal entries of D

the diagonal entries of U and we are done. All of this means that for any invertible matrix A,

$$A = LU = L'D_1D_2U' = L'DU'$$

since the product of two diagonal matrices is a diagonal matrix.

5 Week 5

(1) A point $\mathbf{x} = [x_0 : \cdots : x_n]$ is on a d-dimensional plane if it satisfies

$$a_0x_0 + a_1x_1 + \cdots + a_{d+1}x_{d+1} = 0.$$

We can think of this as a dot product of an x vector and an a vector and an x vector so we can just apply our projective transformation to both sides. If the image of $\mathbf{x}$ is $\mathbf{x}'$, then we have

$$\mathbf{a} \cdot \mathbf{x}' = a_0x_0' + a_1x_1' + \cdots + a_{d+1}x_{d+1}' = 0$$

meaning that the transformation of every point creates a d-dimensional plane.

(2) We can dehomogenize the coordinates by representing $[a : b]$ with a/b for nonzero b and $[a : 0]$ with ∞. This means that we can dehomogenize P_1, P_2, P_3, Q_1, Q_2, Q_3 into $p_1, p_2, p_3, q_1, q_2, q_3 \in k \cup \{\infty\}$. Notice that if there exist a_0, a_1, a_2, a_3 such that

$$\frac{a_0\left(\frac{a}{b}\right) + a_1}{a_2\left(\frac{a}{b}\right) + a_3} = \frac{c}{d},$$

then there exists a projective transformation that maps $[a : b] \to [c : d]$. This is because the matrix for the transformation will just be

$$\begin{pmatrix} a_0 & a_2 \\ a_1 & a_3 \end{pmatrix}.$$

We can verify this by multiplying by $[a : b] = [a/b : 1]$, we get

$$\begin{pmatrix} a/b & 1 \end{pmatrix} \begin{pmatrix} a_0 & a_2 \\ a_1 & a_3 \end{pmatrix} = \begin{pmatrix} a_0 \cdot a/b + a_1 & a_2 \cdot a/b + a_3 \end{pmatrix} \to \left[\frac{a_0\left(\frac{a}{b}\right) + a_1}{a_2\left(\frac{a}{b}\right) + a_3} : 1\right]$$

and this must equal $[c : b] = [c/b : 1]$.

Now this means that we have four variables (the a_i's) and three equations (one for each P_i). This means that there does exist a solution but there also exists a free variable. This is fine, though, because we need to have a unique matrix up to scaling and this definitely holds.

6 Week 6

(1) First we can see that when we do scalar multiplication, we just multiply the coefficients by the scalar. Since F is a field, it is closed so even after scalar multiplication so the vector is still in the set of symmetric polynomials because coefficients will stay the same even after scalar multiplication. Now we need to prove the associative property and the distributive property. Let us start with proving $a \cdot (b \cdot \mathbf{v}) = (ab) \cdot \mathbf{v}$. When we multiply $\mathbf{v}$ by b, we are multiplying the coefficients by b and after this, we multiply the coefficients of the symmetric polynomial by a. This is just the same as multiplying the coefficients by ab and we are done.

Next we move on to the distributive property. When we add two symmetric polynomials, we can think of just adding the coefficients of two like terms. If a term exists in one polynomial and it doesn't in the other, we can think of this just add the missing term but give it a coefficient of 0. This just means that we can think of $\mathbf{u} + \mathbf{v}$ as the sum of two vectors in F^k for some k. However, we know that F^k is a vector space over F so the distributive property must hold.

For the dimension, we know that x_i can have an exponent from 0 to d. This gives us a total of $(d+1)^n$ different terms. Now the terms can be grouped into groups of n since the polynomials are symmetric so the dimension is $\boxed{(d+1)^n/n}$.

(2) First $F[x_1, ..., x_n]^G$ is the field of polynomials that are invariant under G. Again, scalar multiplication just multiplies the coefficients by the scalar so coefficients stay the same so the multiplied vector is still in $F[x_1, ..., x_n]^G$. Now we can check associativity and the distributive property in the same way as the previous problem since still, all we are doing is multiplying the coefficients.

7 Week 7

(1) A strictly diagonally dominant matrix A is a matrix that satifies

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

for all i. For the sake of contradiction, assume A is not invertible. This means that there exists a nonzero $\mathbf{v}$ such that

$$A\mathbf{v} = \mathbf{0}.$$

Now let v_i be the entry with the largest magnitude. This means that

$$\sum_j a_{ij} v_j = 0 \implies a_{ii} v_i = -\sum_{j \neq i} a_{ij} v_j.$$

Therefore,

$$a_{ii} = -\sum_{j \neq i} \frac{v_j}{v_i} a_{ij}$$

so by the Triangle Inequality, we have

$$|a_i i| \leq \sum_{j \neq i} \left| \frac{v_j}{v_i} a_{ij} \right| \leq \sum_{j \neq i} |a_{ij}|$$

which is a contradiction.

For the second part, by the Gershgorin Circle Theorem, the eigenvalues satisfy $\lambda \in D(a_{ii}, \sum_{j \neq i} |a_{ij}|)$ for some i. However, since A is strictly diagonally dominant and since the diagonal entries are all positive, this must mean that the eigenvalues have positive real parts.

(2) There must be a matrix Q such that

$$Q^{-1} A Q = J = \begin{pmatrix} J(\lambda_1) & & & \\ & J(\lambda_2) & & \\ & & \ddots & \\ & & & J(\lambda_n) \end{pmatrix}$$

where

$$J(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

is a $m_i \times m_i$ matrix. Now we define D_i to be the diagonal matrix with diagonal entries $1, t, t^2, ..., t^{m_i}$ for some real t and we define D as the diagonal matrix with diagonal entries $D_1, D_2, ..., D_n$. Now we have

$$D_i^{-1} J(\lambda_i) D_i = \begin{pmatrix} \lambda_i & t & & \\ & \lambda_i & t & \\ & & \ddots & t \\ & & & \lambda_i \end{pmatrix}$$

and

$$D^{-1}JD = D^{-1}Q^{-1}AQD = S^{-1}AS = \begin{pmatrix} D_1^{-1}J(\lambda_1)D_1 & & & \\ & D_2^{-1}J(\lambda_i)D_2 & & \\ & & \ddots & \\ & & & D_n^{-1}J(\lambda_n)D_n \end{pmatrix}$$

where we define $S \equiv QD$. Therefore $G(S^{-1}AS)$ is just the union of disks with centers as the eigenvalues and radii as t or 0 since $S^{-1}AS$ is the matrix with eigenvalues along the diagonal and t along the superdiagonal. Therefore the intersection over all possible t gives us the set of eigenvalues.

(4) First we notice that each disk contains a unique eigenvalue. Now since the matrix is real, the characteristic polynomial of the matrix must have real coefficients. This means that for every root λ, we know that $\bar{\lambda}$ is also a root. In other words, the eigenvalues form complex conjugate pairs. However, the disks are centered on the real axis so the disk contains λ iff it contains $\bar{\lambda}$ so $\lambda = \bar{\lambda}$. Therefore, all the eigenvalues are real.

8 Week 9

(1) Since the rank is 1, all rows are multiples of each other and all columns are multiples of each other. Now when we multiply $\mathbf{u}$ and $\mathbf{v}$, we have a matrix with columns that are multiples of $\mathbf{u}$. We can just make $\mathbf{u}$ a column of B and choose the entries of $\mathbf{v}$ as the right scalar multiples which means we are done.

(2) We know that there exists an A and λ such that

$$A\vec{v} = \lambda\vec{v} = \Delta\vec{v}.$$

However, since λ is the largest eigenvalue, we know that $\vec{v}$ is just the vector of all 1's. This means that A times this vector is a vector of all Δ's so since the entries of A are either 1 or 0, all rows of A must have exactly Δ 1's which means that the graph is Δ-regular.

(3) By Weyl's Theorem, we have

$$\lambda_k(A) + \lambda_1(B) \le \lambda_k(A+B) \le \lambda_k(A) + \lambda_n(B).$$

We can subtract $\lambda_k(A)$ from everything to get

$$\lambda_1(B) \le \lambda_k(A+B) - \lambda_k(A) \le \lambda_n(B).$$

This means that $\lambda_k(A+B) - \lambda_k(A)$ is in between the highest and lowest eigenvalues of B so

$$|\lambda_k(A+B) - \lambda_k(A)| \leq \rho(B).$$