LINEAR ALGEBRA II

Anita Naolekar, notes by Ramdas Singh

Second Semester

List of Symbols

Contents

Chapter 1

PERMUTATION GROUPS

January 3rd.

Let S_n denote the set of all bijections (permutations) on the set $\{1, 2, \ldots, n\}$. If $\sigma, \tau \in S_n$, let us define $\sigma\tau$ to be the bijection defined as

$$(\sigma\tau)(i) = \sigma(\tau(i)) \forall 1 \leq i \leq n. \tag{1.1}$$

This gives us a binary operation on S_n which is associative, and S_n will then contain the identity permutation 1 such that $\sigma 1 = 1\sigma = \sigma$ for all $\sigma \in S_n$. For every such σ, we can also find a $\sigma^{-1} \in S_n$ such that $\sigma\sigma^{-1} = \sigma^{-1}\sigma = 1$. The set S_n equipped with this binary operation, thus, forms a group. In this case, we call S_n as the *symmetric group* of degree n. We now define a cycle in regards to permutations.

Definition 1.1. A *cycle* is a a string of positive integers, say $(i_1, i_2, \ldots, i_k)$, which represents the permutation $\sigma \in S_n$ (with $k \leq n$) such that $\sigma(i_j) = i_{j+1}$ for all $1 \leq j \leq k-1$, and $\sigma(i_k) = i_1$, and fixes all other integers.

We also note that S_3 is the smallest Abelian group possible, upto isomorphism. S_3 is one of the only two groups of order 6, and can be written as

$$S_3 = \{1, \sigma = (1, 2, 3), \sigma^2 = (1, 3, 2), \tau = (1, 2), \sigma\tau = (1, 3), \tau\sigma = (2, 3)\}. \tag{1.2}$$

Some other observations arise. We find that $\sigma^3 = \tau^2 = 1$, and that $\tau\sigma = \sigma^2\tau$. We notice another fact via this σ;

Remark 1.2. A *k-cycle* $\sigma = (i_1, i_2, \ldots, i_k)$ is of order k, that is, $\sigma^k = 1$.

Definition 1.3. Two cycles in S_n are called disjoint if they have no intger in common.

We note that if σ and τ are two disjoint cycles in S_n then σ and τ commute, that is, $\sigma\tau = \tau\sigma$.

Proposition 1.4. *Every σ in S_n can be written uniquely as a product of disjoint cycles.*

Every cycle can be written as a product of 2-cycles. 2-cycles are called *transpositions*. This can easily be seen as

$$(a_1, a_2, \ldots, a_n) = (a_1, a_n)(a_1, a_{n-1}) \cdots (a_1, a_3)(a_1, a_2). \tag{1.3}$$

1.1 Even and Odd Permutations

Let $x_1, x_2, \ldots, x_n$ be indeterminates, and let

$$\Delta = \prod_{1 \leq i < j \leq n} (x_i - x_j). \tag{1.4}$$

Let $\sigma \in S_n$, and define

$$\sigma(\Delta) = \prod_{1 \leq i < j \leq n} (x_{\sigma(i)} - x_{\sigma(j)}). \tag{1.5}$$

We find that $\sigma(\Delta) = \pm\Delta$. Based on this, we classify permutations as odd or even.

Definition 1.5. A permutation σ is said to be an *even permutation* if $\sigma(\Delta) = \Delta$, and is said to be an *odd permutation* if $\sigma(\Delta) = -\Delta$. The sign of a permutation σ, denoted by $\epsilon(\sigma)$, is $+1$ if σ is even, and is -1 if σ is odd. So, $\sigma(\Delta) = \epsilon(\sigma)\Delta$.

Proposition 1.6. *The map $\epsilon : S_n \to \{-1, +1\}$, where $\epsilon(\sigma)$ is the sign of σ, is a homomorphism, that is, $\epsilon(\sigma\tau) = \epsilon(\sigma)\epsilon(\tau)$ for all $\sigma, \tau \in S_n$.*

Proof. Start with $\tau(\Delta)$;

$$\tau(\Delta) = \prod_{1 \le i < j \le n} (x_{\tau(i)} - x_{\tau(j)}). \tag{1.6}$$

Let there be k factors of this polynomial where $\tau(i) > \tau(j)$ with $i < j$. We find that $\tau(\Delta) = (-1)^k\Delta$, and so, $\epsilon(\tau) = (-1)^k$. Now, $\sigma\tau(\Delta)$ has exactly k factors of the form $x_{\sigma(j)} - x_{\sigma(i)}$, with $j > i$. Bringing out a factor $(-1)^k$, we find that $\sigma\tau(\Delta)$ has all factors of the form $x_{\sigma(i)} - x_{\sigma(j)}$, with $j > i$. Thus,

$$\epsilon(\sigma\tau)\Delta = \sigma\tau(\Delta) = (-1)^k \prod_{1 \le i < j \le n} (x_{\sigma(i)} - x_{\sigma(j)}) = (-1)^k\sigma(\Delta) = (-1)^k\epsilon(\sigma)\Delta = \epsilon(\tau)\epsilon(\sigma)\Delta. \tag{1.7}$$

Cancelling out the Δ, we find $\epsilon(\sigma\tau) = \epsilon(\sigma)\epsilon(\tau)$. ■

ϵ is a homomorphism to an Abelian group, so $\epsilon(\sigma\tau) = \epsilon(\sigma)\epsilon(\tau) = \epsilon(\tau)\epsilon(\sigma)$.

Proposition 1.7. *If $\lambda = (i, j)$ is a transposition, then $\epsilon(\lambda) = -1$.*

Proof. If $\lambda = (1, 2) \in S_n$, it is easy to show that

$$\lambda(\Delta) = (x_1 - x_2) \cdots (x_1 - x_n)(x_2 - x_3) \cdots (x_2 - x_n) \cdots = (-1)(\Delta). \tag{1.8}$$

Now, if $\sigma = (i, j)$, with $(i, j) \ne (1, 2)$, then $(i, j) = \lambda(1, 2)\lambda$ where λ interchanges 1 and i, and interchanges 2 and j. Using that fact that ϵ is a homomorphism, $\epsilon(\sigma) = -1$. ■

A cycle σ of length k is an even permutation if and only if k is odd. This is because it can be decomposed into $k - 1$ transpositions, and we would then have $\epsilon(\sigma) = (-1)^{k-1} = 1$ (using the fact that ϵ is a homomorphism). Some more corollaries of the previous proposition include the fact that ϵ is a surjective map, and that $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$.

If, for $\sigma \in S_n$, σ can be decomposed as $\sigma_1\sigma_2 \cdots \sigma_k$, where σ_i is a m_i-cycle, then $\epsilon(\sigma_i) = (-1)^{m_i - 1}$, and $\epsilon(\sigma) = (-1)^{(\sum m_i) - k}$.

Proposition 1.8. *σ is an odd permutation if and only if the number of cycles of even length in its cycle decomposition is odd.*

1.2 The Determinant

Definition 1.9. If $A = (a_{ij})$ is a square matrix of order n, then the *determinant* of A is defined as

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{1.9}$$

Using this definition of the determinant of a square matrix, one may derive the usual determinant properties with ease.

January 7th.

Remark 1.10. The following properties may be inferred:

- If A contains a row of zeroes, or a column of zeroes, then $\det A = 0$.
- $\det I_n = 1$.
- The determinant of a diagonal matrix is the product of the diagonal elements. This is because if $\sigma \in S_N$ is not the identity permutation, then there exists at least one element in the corresponding term where $i \ne \sigma(i)$, and $a_{i\sigma(i)}$ makes the term zero. For the identity transformation, it contains only those elements of the form a_{ii}.

Other non-trivial properties may also be shown with ease.

Corollary 1.11. *If A is an upper triangular matrix, then $\det A$ is the product of the diagonal entries.*

Proof. If $a_{1\sigma(1)} \cdots a_{n\sigma(n)} \neq 0$, then $a_{n\sigma(n)} \neq 0$, that is, $\sigma(n) = n$, as $a_{ni} = 0 \ \forall \ i < n$. Again, $\sigma_{(n-1)\sigma(n-1)} \neq 0$ leads us to conclude that $\sigma(n-1) = n-1$ as σ is a bijection and has to lead to a non-zero element. By similar logic, $\sigma(i) = i$ for all valid i. So, σ is the identity permutation. ■

Corollary 1.12. *If A is a lower triangular matrix, then $\det A$ is the product of the diagonal entries.*

Proof. The proof of this is similar to the previous proof if we consider that the determinant of the tranpose of a matrix is equal to the determinant of said matrix. ■

Theorem 1.13. *The determinant of a matrix is equal to the determinant of its transpose, that is, $\det A = \det A^t$ for a square matrix A.*

Proof. The proof is left as an exercise to the reader. ■

Proposition 1.14. *Let B be obtained from A by multiplying a row (or column) of A by a non-zero scalar, α. Then, $\det B = \alpha \det A$.*

Proof. The proof is left as an exercise to the reader. ■

Proposition 1.15. *If B is obtained from A by interchanging any two rows (or columns) of A, then $\det B = -\det A$.*

Proof. Let B be obtained from A by interchanging the rows k and l, with $k < l$. We then have

$$\begin{aligned} \det B &= \sum_{\sigma \in S_n} \epsilon(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{(k-1)\sigma(k-1)} a_{l\sigma(k)} \sigma_{(k+1)\sigma(k+1)} \cdots a_{k\sigma(l)} \cdots a_{n\sigma(n)}. \end{aligned} \tag{1.10}$$

As σ runs through all elements in S_n, $\tau = \sigma(k,l)$ also runs through all S_n. Hence, via $\epsilon(\tau) = -\epsilon(\sigma)$, the equation now looks like

$$\det B = -\sum_{\tau \in S_n} \epsilon(\tau) a_{1\tau(1)} \cdots a_{l\tau(l)} \cdots a_{k\tau(k)} \cdots a_{n\tau(n)} = -\det A. \tag{1.11}$$

■

Proposition 1.16. *If two rows (or columns) of A are equal, then $\det A = 0$.*

Proof. Suppose that the rows k and l of A are equal. Interchanging will alter the determinant by -1, so $\det A = -\det A \implies 2\det A = 0 \implies \det A = 0$ if $2 \neq 0$ in the field F from where the elements of A arrive.

If $2 = 0$ in F, that is, F is of characteristic 2, we pair the σ term in the expression of $\det A$ with the term τ where $\tau = \sigma(k,l)$. The terms corresponding to σ and τ in the expressions are the same, differing in only the sign. Hence, $\det A = 0$. ■

Theorem 1.17. *For a fixed k, let the row k of A be the sum of the two row vectors X^t and Y^t, that is, $a_{kj} = x_j + y_j$ for all $1 \leq j \leq n$. Then $\det A = \det B + \det C$ where B is obtained from A by replacing the row k of A by the row vector X^t, and C is obtained from A by replacing the row k of A by the row vector Y^t.*

Proof. We utilize the fact that $a_{kj} = x_j + y_j$. We have

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{k\sigma(k)} \cdots a_{n\sigma(n)} \\ &= \left(\sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots x_{\sigma(k)} \cdots a_{n\sigma(n)} \right) + \left(\sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots y_{\sigma(k)} \cdots a_{n\sigma(n)} \right) \\ &= \det B + \det C. \end{aligned}$$

■

Proposition 1.18. *If a scalar multiple of a row (or column) is added to a row (or column) of a matrix, the determinant remains unchanged.*

Proof. The proof follows immediately from the previously proved properties. ■

January 10th.

Definition 1.19. For $a_{ij} \in A$, the *cofactor* of a_{ij} is $A_{ij} = (-1)^{i+j} \det M_{ij}$, where M_{ij} is the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i^{th} row and j^{th} column of A.

Lemma 1.20. *Fix k, j. If $a_{kl} = 0$ for all $l \neq j$, then* $\det A = a_{kj} A_{kj}$.

Proof. Take A to be a $n \times n$ matrix. We deal in cases.

- Case I: $k = j = n$. In the expansion of the determinant,

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)},$$

  only those σ's survive where $\sigma(n) = n$. These σ's can be thought of as permutations of S_{n-1} instead. The sign of $\sigma \in S_n$ and $\sigma \in S_{n-1}$ is the same as n is fixed. Thus, we get

$$a_{nn} \sum_{\sigma \in S_{n-1}} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{(n-1)\sigma(n-1)} = a_{nn} \det M_{nn} = (-1)^{n+n} a_{nn} A_{nn} = a_{nn} A_{nn}. \tag{1.12}$$

- Case II: $(k, j) \neq (n, n)$. We construct a matrix B by interchanging $n-k$ rows and $n-j$ columns to bring a_{ij} to the position (n, n). Thus, we have $\det B = (-1)^{n-k+n-j} \det A = (-1)^{k+j} \det A$. But $B = a_{kj} \det M_{kj}$, so

$$\det A = (-1)^{k+j} a_{kj} \det M_{kj} = a_{kj} A_{kj}. \tag{1.13}$$

■

Theorem 1.21. *Let A be a $n \times n$ matrix, and let $1 \leq k \leq n$. Then,* $\det A = \sum_{j=1}^{n} a_{kj} A_{kj}$, *expansion by the k^{th} row.*

Proof. Write out the k^{th} row of A as $x_1^t + \ldots + x_n^t$, where $x_i = (0, \ldots, 0, a_{ki}, 0, \ldots, 0)^t$, and all the other rows remaining are the same. Writing the matrix A as the sum of n matrices where each matrix is the same as A but with a row that looks like x_i^t, we can easily show that $\det A = \sum_{j=1}^{n} a_{kj} A_{kj}$. ■

Example 1.22. Let $n \geq 1$, and let $A_n = \begin{pmatrix} a_1^{n-1} & a_1^{n-2} & \ldots & a_1 & 1 \\ a_2^{n-1} & a_2^{n-2} & \ldots & a_2 & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_n^{n-1} & a_n^{n-2} & \ldots & a_n & 1 \end{pmatrix}$. Then, $\det A_n = \prod_{1 \leq i \leq j \leq n} (a_i - a_j)$.

Proof. If $a_i = a_j$ for some $i \neq j$, then $\det A_n = 0$ as two rows are then identical. Hence, assume that the a_i's are distinct. Now construct

$$B_n = \begin{pmatrix} x_1^{n-1} & x_1^{n-2} & \ldots & x_1 & 1 \\ a_2^{n-1} & a_2^{n-2} & \ldots & a_2 & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_n^{n-1} & a_n^{n-2} & \ldots & a_n & 1 \end{pmatrix}. \tag{1.14}$$

Notice that $\det B_n \in F[x]$, where F is the field, and x is an indeterminate. $\det B$ is also of degree $(n-1)$; let us call this polynomial $f(x)$. Each of $a_2, \ldots, a_n$ are roots of $f(x)$, so $f(x)$ must be of the form $f(x) = C(x - a_2) \ldots (x - a_n)$. Equating coefficients of x^{n-1}, we get

$$C = \prod_{2 \leq i < j \leq n} (a_i - a_j) = \det \begin{pmatrix} a_2^{n-2} & \ldots & a_2 & 1 \\ \ldots & \ldots & \ldots & \ldots \\ a_n^{n-2} & \ldots & a_n & a_1 \end{pmatrix}. \tag{1.15}$$

Thus, we must have

$$f(x) = \left(\prod_{2 \le i < j \le n} (a_i - a_j) \right) (x - a_2) \cdots (x - a_n) \tag{1.16}$$

$$\implies \det A_n = f(1) = \prod_{1 \le i < j \le n} (a_i - a_j). \tag{1.17}$$

■

Example 1.23. Show that there exists a unique polynomial of degree n that takes arbitrary prescribed values at the $(n+1)$ points $x_0, x_1, \ldots, x_n$.

Chapter 2

LINEAR TRANSFORMATIONS

2.1 A Brief Summary

Let $\mathcal{B} = (v_1, \ldots, v_n)$ be a basis of vector space V and $\mathcal{C} = (w_1, \ldots, w_n)$ be a basis of a vector space W. As these are bases, given a $v \in V$, there exists a unique $X \in F^n$ such that $v = \mathcal{B}X$, called the *coordinate vector* of v with respect to the basis $\mathcal{B}$. We note that since the mapping from a $v \in V$ to a $X \in F^n$ is linear in nature and is bijection, the vector spaces V and F^n are isomorphic to each other. Similarly, a mapping that takes $w \in W$ to $Y \in F^m$ shows that W and F^m are isomorphic to each other.

Now suppose that there exists a linear map that takes $v \mapsto Tv$ with $v \in V$ and $Tv \in W$. This transformer T is with respect to the bases $\mathcal{B}$ and $\mathcal{C}$ of V and W, respectively. We construct the $m \times n$ matrix A so that the j^{th} column of A is the coordinate vector of Tv_j with respect to the basis $\mathcal{C}$. We will then have $T(\mathcal{B}) = \mathcal{C}A$. For any vector $v \in V$, we have

$$\begin{aligned} v &= \mathcal{B}X = v_1x_1 + \ldots v_nx_n \\ \implies T(v) &= T(v_1)x_1 + \ldots T(v_n)x_n = (T(v_1), \ldots, T(v_n)) \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = T(\mathcal{B})X = (\mathcal{C}A)X \qquad (2.1) \\ &= (w_1, \ldots, w_m)AX; \qquad (2.2) \end{aligned}$$

the coordinate vector of Tv with respect to the basis AX. In fact, if we denote the isomorphism from V to F^n by $\phi_{\mathcal{B}}$ and the isomorphism from W to F^m by $\phi_{\mathcal{C}}$, we get $\phi_{\mathcal{C}} \circ T = (\text{mult. by } A) \circ \phi_{\mathcal{B}}$.

The next theorem will be divided into two parts.

Theorem 2.1. 1. *The vector space form. Let $T : V \to W$ be a linear mapping between finite dimensional vector spaces V and W, of dimensions n and m respectively. There are bases $\mathcal{B}$ and $\mathcal{C}$ of V and W respectively such that the matrix of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$ looks like* $\begin{pmatrix} I_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{pmatrix}_{m\times n}$.

2. *The matrix form. If A is a $m \times n$ matrix, then there exists an invertible matrix $Q_{m\times m}$ and an invertible matrix $P_{n\times n}$ such that $Q^{-1}AP$ is of the form $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, where r is the rank of A.*

3. *In fact, both these forms of the theorem are equivalent.*

Proof. 1. Let $(u_1, \ldots, u_{n-r})$ be a basis of $\ker T$. We can extend this to a basis $\mathcal{B}$ by appending independent vectors that do not belong to the kernel of T, that is, $(v_1, \ldots, v_r, u_1, \ldots, u_{n-r})$. Let $(Tv_1, \ldots, Tv_r)$ be a basis of $\mathrm{Im}T$. We can extend this to a basis of W, say $\mathcal{C} = (w_1, \ldots, w_r, w_{r+1}, \ldots, w_m)$, where $w_i = Tv_i$ for $1 \leq i \leq r$. These bases are the desired ones.

2. P is a sequence of column operations, multipled to form a matrix, and Q^{-1} is a sequence of row operations, multiplied to form a matrix, that get the matrix A into the desired form. These are our desired P and Q.

3. Suppose the vector space form holds. Let A be a $m \times n$ matrix over F, with $A : F^n \to F^m$ defined as $X \mapsto AX$. There then exists a basis $\mathcal{B}$ of F^n and a basis $\mathcal{C}$ of F^m such that the linear map A with respect to ther bases $\mathcal{B}$ and $\mathcal{C}$ has the desired matrix. We then have $\mathcal{B} = I_n P_{n\times n}$ and $\mathcal{C} = I_m Q_{m\times m}$, with both P and Q invertible. We claim that the matrix of the linear mapping A with respect to the bases $\mathcal{B}$ and $\mathcal{C}$ is $Q^{-1}AP$.

■

Appendices

Chapter A

Appendix

Extra content goes here.

Appendix

Index