LINEAR ALGEBRA II

Anita Naolekar, notes by Ramdas Singh

Second Semester

List of Symbols

Contents

1 PERMUTATION GROUPS 1
- 1.1 Even and Odd Permutations . . . 1
- 1.2 The Determinant . . . 2

2 EIGENVECTORS AND EIGENVALUES 7
- 2.1 Linear Transformers and an Introduction . . . 7
  - 2.1.1 Linear Operators . . . 8
  - 2.1.2 Eigenvectors and Eigenvalues . . . 8
- 2.2 Finding Eigenvalues and Eigenvectors . . . 9
  - 2.2.1 Eigenspace . . . 10
- 2.3 Diagonalizability . . . 10
- 2.4 Polynomials . . . 11
  - 2.4.1 Interaction with Linear Operators . . . 11

Appendices 15

A Appendix 17

Index 19

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 $(i, j) \neq (1, 2)$, 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 $i \neq \sigma(i)$, 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, $\det A = \sum\limits_{j=1}^{n} a_{kj} A_{kj}$, 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, $\det A_n = \prod\limits_{1 \leq i \leq j \leq n} (a_i - a_j)$.

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

EIGENVECTORS AND EIGENVALUES

2.1 Linear Transformers and an Introduction

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

$$v = \mathcal{B}X = v_1x_1 + \ldots v_nx_n$$

$$\implies T(v) = T(v_1)x_1 + \ldots T(v_n)x_n = (T(v_1), \ldots, T(v_n)) \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = T(\mathcal{B})X = (\mathcal{C}A)X \tag{2.1}$$

$$= (w_1, \ldots, w_m)AX; \tag{2.2}$$

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 $1 \le i \le r$. 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 $Q^{-1}AP$.

■

January 16th.

Proposition 2.2. *1. Let $T : V \to W$ be a linear map, and A the matrix of T with respect to the bases $\mathcal{C}$ and $\mathcal{C}$ of V and W respectively. Let $\mathcal{B}'$ and $\mathcal{C}'$ be new bases of V and W respectively, and let the change of basis matrices be given by $\mathcal{B}' = \mathcal{B}P$ and $\mathcal{C}' = \mathcal{C}Q$. Then the matrix of T with respect to $\mathcal{B}'$ and $\mathcal{C}'$ is $Q^{-1}AP$.*

2. If $A' = Q_1^{-1}AP_1$, where P_1 and Q_1 are $n \times n$ and $m \times m$ invertible matrices, respectively, then A' is the matrix of T with respect to the bases $\mathcal{B}P_1$ and $\mathcal{C}Q_1$.

Proof. Let the coordinate vector of v with respect to the basis $\mathcal{B}'$ be X'. We claim that the coordinate vector of Tv with respect to the basis $\mathcal{C}'$ is Y', where $Y' = (Q^{-1}AP)X'$. We assume that $\mathcal{B}' = \mathcal{B}P_{n\times n}$, $\mathcal{C}' = \mathcal{C}Q_{m\times m}$, and $T(\mathcal{B}) = \mathcal{C}A_{m\times n}$. If $v = \mathcal{B}X$, then $T(v) = \mathcal{C}(AX)$. If we let $v = \mathcal{B}'\mathcal{X}' = v_1'x_1' + \ldots + v_n'x_n'$, then

$$T(v) = \mathcal{C}'Y' = (\mathcal{C}Q)' = \mathcal{C}(QY') = \mathcal{C}(APX') \implies QY' = APX' \implies Y' = (Q^{-1}AP)X' \tag{2.3}$$

To prove the second part, we will show that the first part implies it. Let $A_{m\times n}$ be a matrix. Let T_A be the linear map from $\mathbb{R}^n \to \mathbb{R}^m$ given by multiplication by A, that is $T_A : \mathbb{R}^n \to \mathbb{R}^m$ given by $X \mapsto AX$. By the first part, there exist bases $P_{n\times n}$ and $Q_{m\times m}$, both invertible, such that with respect P and Q, the matrix of T_A looks like $\begin{pmatrix} I & O \\ O & O \end{pmatrix}$, that is, $Q^{-1}AP = \begin{pmatrix} I & O \\ O & O \end{pmatrix}$. ■

2.1.1 Linear Operators

Let $T : V_{\mathcal{B}} \to V_{\mathcal{B}}$. Let A be the matrix of T with respect to the basis $\mathcal{B}$. The other matrices of T with respect to new bases are $P^{-1}AP$, where $P_{n\times n}$ is invertible. Also, the fact that T is bijective, one-one, or onto are all equivalent for a finite dimensional vector space V.

2.1.2 Eigenvectors and Eigenvalues

Definition 2.3. A non-zero vector $v \in V$ is said to be an *eigenvector* of T if $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$. If A is a $n \times n$ matrix, a non-zero column vector X is said to be an eigenvector of A if $AX = \lambda X$ for some $\lambda \in \mathbb{F}$. λ, in both these cases, is called the *eigenvalue* of v and X respectively.

Usually, we always disregard the zero vector being an eigenvector. If v is an eigenvector of $T : V \to V$, and $v = \mathcal{B}X$ with respect to some basis $\mathcal{B}$ of V, then X is an eigenvector of the matrix of T with respect to the basis $\mathcal{B}$. In fact,

$$\mathcal{B}(AX) = (\mathcal{B}A)X = T(\mathcal{B})X = T(\mathcal{B}X) = Tv = \lambda v = \lambda\mathcal{B}X = \mathcal{B}(\lambda X) \implies AX = \lambda X. \tag{2.4}$$

The converse is also true; if X is an eigenvector of $A_{n\times n}$, then X is also an eigenvector of $T_A : \mathbb{R}^n \to \mathbb{R}^n$.

Proposition 2.4. *0 is an eigenvalue of $A_{n\times n}$ ($T : V \to V$) if and only if A (T) is non-invertible (not an isomorphism).*

Suppose v is an eigenvector of $T : V \to V$ with eigenvalue λ. Let W be the subspace spanned by v. Then every vector $w \in W$ is an eigenvector of T with eigenvalue λ. The proof of this is left as an exercise.

Definition 2.5. Two matrices $A'_{n\times n}$ and $A_{n\times n}$ are called *similar matrices* if there exists an invertible matrix $P_{n\times n}$ such that $P^{-1}AP = A'$.

Again let $T : V \to V$ be a linear operator, and let $\mathcal{B} = (v_1, \dots, v_n)$. Suppose, with respect to the basis $\mathcal{B}$, the matrix of T is $\begin{pmatrix} \lambda_1 & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots \end{pmatrix}$. Then v_1 is an eigenvector with eigenvalue λ_1.

2.2 Finding Eigenvalues and Eigenvectors

January 21st.

Let $T : V \to V$ and let $\mathcal{B} = (v_1, \dots, v_n)$ be a basis of V. Then the matrix of T with respect to the basis $\mathcal{B}$ is a diagonal matrix if and only if each of the basis elements is an eigenvector. An equivalent statement for matrices is that an $n \times n$ matrix A is similar to a diagonal matrix if and only if $\mathbb{F}^n$ admits a basis consisting of eigenvectors of A. The proof of this is left as an exercise to the reader.

We can now discuss the computation. For a linear operator $T : V \to V$, λ is an eigenvalue of T if and only if there exists a non-zero vector v such that $Tv = \lambda v$. This can be rearranged to give

$$(\lambda I_v - T)v = 0. \tag{2.5}$$

We can now consider $\lambda I_v - T : V \to V$ to be a linear operator which maps $v \mapsto \lambda v - Tv$. If eigenvalues exist, this operator is a singular operator, that is, it contains a non-trivial kernel. The matrix of the operator $\lambda I_v - T$ comes out to be $\lambda I_n - A$, where A is the matrix of T with respect to the basis $\mathcal{B}$. This matrix is now singular, so we must have

$$\det(\lambda I_n - A) = 0. \tag{2.6}$$

The equation $\det(\lambda I_n - A)$ is called the *characteristic polynomial* of A, and also $T(?)$. The roots of this polynomial in λ which lie in $\mathbb{F}$ are the eigenvalues of A, and T as well.

We would now like to show that similar matrices have the same eigenvalues, that is,

$$\det(\lambda I_n - P^{-1}AP) = \det(\lambda I_n - A). \tag{2.7}$$

This is simple to see as $\det(\lambda I_n - P^{-1}AP) = \det(P^{-1}(\lambda I_n - A)P) = \det P^{-1} \cdot \det(\lambda I_n - A) \cdot \det P = \det(\lambda I_n - A)$. The found out eigenvalues from this equation can then be put back and solved for v to get the corresponding eigenvectors.

Proposition 2.6. *Let $\lambda_1, \dots, \lambda_r$ be distinct eigenvalues of $T : V \to V$ and let $v_1, \dots, v_r$ be the corresponding eigenvectors of T. Then $(v_1, \dots, v_r)$ is a linearly independent set in V.*

Proof. We claim that this is true for $r = 1, 2$. Using a form of induciton, we will assume the result for $r - 1$. Begin with

$$\begin{aligned} \alpha_1 v_1 + \ldots + \alpha_r v_r &= 0 \\ \implies \alpha_1 T v_1 + \ldots + \alpha_r T v_r &= 0 \\ \implies \alpha_1 \lambda_1 v_1 + \ldots + \alpha_r \lambda_r v_r &= 0. \end{aligned} \tag{2.8}$$

Multiplying the first equation by λ_1 and subtracting it from the current equation, we have

$$\begin{aligned} (\alpha_2\lambda_2 - \alpha_2\lambda_1)v_2 + (\alpha_3\lambda_3 - \alpha_3\lambda_1)v_3 + \ldots + (\alpha_r\lambda_r - \alpha_r\lambda_1)v_r &= 0 \\ \implies \alpha_2(\lambda_2 - \lambda_1) + \alpha - 3(\lambda_3 - \lambda_1)v_3 + \ldots + \alpha_r(\lambda_r - \lambda_1)v_r &= 0. \end{aligned} \tag{2.9}$$

By hypothesis, $\alpha_j(\lambda_j - \lambda_1) = 0$. As the eigenvalues are distinct, we must have $\alpha_j = 0$ for $j = 2, 3, \dots, r$. We are left with $\alpha_1 v_1 = 0$, which gives us $\alpha_1 = 0$. ■

When the n eigenvalues found of A are distinct, the corresponding eigenvectors $v_1, \dots, v_n$ are linearly independent in $\mathbb{F}^n$, and hence $\mathcal{B} = (v_1, \dots, v_n)$ is a basis of $\mathbb{F}^n$. The matrix $P^{-1}AP$ is the matrix of the linear operator $T_A : \mathbb{F}^n \to \mathbb{F}^n$ with respect to the basis $\mathcal{B}$, with the column of P being the eigenvectors $v_1, \dots, v_n$. As $\mathcal{B}$ consists of only eigenvectors, $P^{-1}AP$ is a diagonal matrix with the diagonal entries being the n eigenvalues.

We now define the determinant and trace for a linear operator. For such an operator T, $\mathrm{tr}T = \mathrm{tr}A$ where A is a matrix of T with respect to some abitrary basis. Note that since $\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(APP^{-1}) = \mathrm{tr}A$, the choice of basis is not important. Similarly, we define $\det T = \det A$.

We can now have a closer look at the characteristic equation. To find the constant term of $\det(xI - A)$, we simply plug in $x = 0$ to give us $\det(-A) = (-1)^n \det A$. The coefficient of x^{n-1} in $\det(xI - A)$ is $-\mathrm{tr}A$ as the coefficients of x^{n-1} come solely from the expansion of $(x - a_{11})(x - a_{22}) \cdots (x - a_{nn})$. Clearly, we can conclude that the sum of the eigenvalues is $\mathrm{tr}A$ and the product of the eigenvalues is $\det A$.

2.2.1 Eigenspace

January 23rd.

For ease, let us denote $\chi_T(x)$ to mean $\det(xI - A)$. The *eigenspace* for a given eigenvalue λ is defined as

$$E_\lambda = \{v \in V : Tv = \lambda v\}. \tag{2.10}$$

This is a subspace of the vector space V. The *geometric multiplicity* of λ is defined as the dimension of E_λ. This geometric multiplicity of λ is always less than or equal to its algebraic multiplicity in $\chi_T(x)$. For recall, the *algebraic multiplicity* of λ is the highest power of $(x-\lambda)$ that divides $\chi_T(x)$.

Theorem 2.7. *Let λ be an eigenvalue of $T : V \to V$. Then the geometric multiplicity of λ is always less than or equal to its algebraic multiplicity.*

Proof. Let k me the geometric multiplicity of λ. Let $(v_1, \ldots, v_k)$ be an ordered basis of E_λ. Extend this to a basis $\mathcal{B} = (v_1, \ldots, v_k, u_1, \ldots, u_{n-k})$ of V. The matrix of T with respect to the basis $\mathcal{B}$ is of the form $A = \begin{pmatrix} \lambda I_k & B \\ O & D \end{pmatrix}$. Thus, the characteristic polynomial looks like

$$\chi_T(x) = \det(xI_n - A) = \det \begin{pmatrix} (x-\lambda)I_k & -B \\ O & xI_{n-k} - D \end{pmatrix} = (x-\lambda)^k \cdot \det(XI_{n-k} - D). \tag{2.11}$$

This shows that $(x-\lambda)^k$ divides $\chi_T(x)$, so we must have an algebraic multiplicity greater than or equal to this k. ■

2.3 Diagonalizability

We first define what this means for a linear mapping from V to V.

Definition 2.8. A linear operator $T : V \to V$ is said to be a *diagonizable linear operator* if there exists a basis of V consisting of eigenvectors of T. This means that the matrix of T with respect to this basis if a digaonal matrix and the matrix of T with respect to any other basis is similar to this diagonal matrix.

A similar definition works for matrices.

Definition 2.9. An $n \times n$ matrix A over $\mathbb{F}$ is said to be a *diagonizable matrix* if A is similar to a diagonal matrix. Equivalently, $\mathbb{F}^n$ then admits a basis consisting of eigenvectors of A, thinking of $T_A : \mathbb{F}^n \to \mathbb{F}^n$ as a linear operator.

Now let us suppose that T is diagonizable. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. There then exists an ordered basis consisting of eigenvectors of T and with respect to this basis, the matrix of T is a diagonal matrix with diagonal entries consisting solely of $\lambda_1, \lambda_2, \ldots, \lambda_k$.

If λ_i is of algebraic multiplicity d_i, then the matrix of T looks like $\begin{pmatrix} \lambda_1 I_{d_1} & & & \\ & \lambda_2 I_{d_2} & & \\ & & \cdots & \\ & & & \lambda_k I_{d_k} \end{pmatrix}$. Thus, the characteristic polynomial then looks like $(x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\cdots(x-\lambda_k)^{d_k}$.

The geometric multiplicity of λ_i is the dimension of E_{λ_i}, that is, the nullity of the operator $(\lambda_i I_n - A)$. But here, $\ker(\lambda_i I_n - A) = d_i$, which is just the algebraic multiplicity of λ_i. Hence, if T is diagonizable, then each eigenvalue of it has the same algebraic multiplicity and geometric multiplicity.

Proposition 2.10. *If $E_{\lambda_1}, \ldots, E_{\lambda_k}$ are the eigenspaces corresponding to the distinct eigenvalues, say, $\lambda_1, \ldots, \lambda_k$ of T, then $E = E_{\lambda_1} + \cdots + E_{\lambda_k}$ is a direct sum.*

Proof. It is enough to show that $E_{\lambda_1}, \ldots, E_{\lambda_k}$ are independent. Let $v_1 + v_2 + \ldots + v_k = 0$, where $v_i \in E_{\lambda_i}$. As $v_1, v_2, \ldots, v_k$ come from distinct eigenspaces, they are linearly independent, and our equation must imply that $v_1 = \ldots = v_k = 0$. ■

Proposition 2.11. *If T is a diagonizable operator, and if $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T, then*

$$V = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_k}. \tag{2.12}$$

Proof. As T is diagonizable, the algebraic and geometric multiplicities are equal for all the eigenvalues λ_i. Denote $\dim E_{\lambda_i} = d_i$. As $\chi_T(x)$ completely factors into linear factors, due to T being diagonizable, we have $n = d_1 + \ldots + d_k$. Also, $E_{\lambda_1} + \ldots + E_{\lambda_k}$ is a direct sum, that is,

$$\dim(E_{\lambda_1} + \ldots + E_{\lambda_k}) = \dim E_{\lambda_1} + \ldots + \dim E_{\lambda_k} = n. \tag{2.13}$$

This direct sum is a subspace of V and has the dimension as V. This mut mean that the direct sum is exactly V. ■

Theorem 2.12. *Let T be a linear operator on a finite dimensional vector space V, and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. Also let E_{λ_i} be the eigenspace of λ_i. Then, the following are equivalent.*

- *T is diagonizable,*
- *$\chi_T(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_k)^{d_k}$ and $\dim E_{\lambda_i} = d_i$,*
- *$V = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_k}$.*

2.4 Polynomials

January 28th.

Let $\mathbb{F}[x]$ denote the set of all polynomials with coefficients coming from the field $\mathbb{F}$. With respect to the addition, it is an Abelian group. The multiplication here is associative, commutative, and distributive; there also exists a multiplicative identity. This makes $\mathbb{F}[x]$ into a commutative ring. Note that $\mathbb{F}[x]$ is also an infinite dimensional vector space over $\mathbb{F}$, since scalar multiplication is also defined. Together, these combine to form an algebra over the field.

Definition 2.13. Let $d \in \mathbb{F}[x]$ with $d \neq 0$. For $f \in \mathbb{F}[x]$, we say that d divides f if there exists a $q \in \mathbb{F}[x]$ such that $f = dq$ in $\mathbb{F}[x]$.

Corollary 2.14. *For $f \in \mathbb{F}[x]$, $f(c) = 0$ if and only if $x - c$ divides $f(x)$.*

Corollary 2.15. *A polynomial $f \in \mathbb{F}[x]$ of degree n has at most n roots in $\mathbb{F}$.*

Proof. The proof is by induction. Note that this is true for $n = 0, 1$. If α is a root, then $f(x) = (x-\alpha)\cdot q(x)$. As $q(x)$ is of degree $n - 1$, and all roots of $q(x)$ are root of $f(x)$, this follows by hypothesis. ■

Definition 2.16. An *ideal* of $\mathbb{F}[x]$ is a subspace of $\mathbb{F}[x]$, say I, such that if $f \in I$ and $g \in \mathbb{F}[x]$, then $fg \in I$.

Example 2.17. Let $f \in \mathbb{F}[x]$. Define $I_f = \langle f \rangle = \{fg : g \in \mathbb{F}[x]\}$. Note that I_f is called a *principal ideal*, that is, it is an ideal generated by a single element.

Theorem 2.18. *$\mathbb{F}[x]$ is a principal ideal domain, that is, every ideal in $\mathbb{F}[x]$ is a principal ideal.*

Proof. Let d be a polynomial of least degree in the ideal I, where I is a non-zero ideal. Let, without loss of generatlity, d be monic (if not, simply multiply it by a sutitable scalar).

Let $f \in I$. Then there exists $q, r \in \mathbb{F}[x]$ such that $f = dq + r$ and either $r = 0$ or $\deg r < \deg d$. Note that since $f, d \in I$, $dq \in I$, so $f - dq \in I \implies r \in I$. As d was of minimal degree in I, we must have $r = 0$. Thus, $f = dq$ and, thus, $I = \langle d \rangle$. ■

If I is an ideal of $\mathbb{F}[x]$, then there exists a unique polynomial $d \in I$ such that $I = \langle d \rangle$.

2.4.1 Interaction with Linear Operators

Let $f \in \mathbb{F}[x]$, and let $T : V \to V$ be a linear mapping. If

$$f(x) = a_0 + a_1 x + \ldots + a_k x^k$$

with $a_k \neq 0$, we define

$$f(T) = a_0 I_n + a_1 T + \ldots + a_k T^k.$$

Note that $f(T)$ is also a linear mapping from V to V. Let I be the set of all $f \in \mathbb{F}[x]$ such that $f(T)$ is the zero operator. All such polynomials are called *annihilator*s. I satisfies the properties of a vector space; it is a subspace of the space of all polynomials. I is also an ideal of $\mathbb{F}[x]$.

Definition 2.19. The *minimal polynomial* of the linear operator $T : V \to V$ is the generator of the ideal of annihilators.

Denote the minimal polynomial by $m_T(x)$. So, $m_T(x)$ is

1. monic,
2. of least degree among all annihilators of T.

If A is a $n \times n$ matrix, the minimal polynomial of A is defined as the unique monic polynomial $m_A(x)$ of least degree such that $m_A(A) = O_{n\times n}$. It can be verified that if A is the matrix of a linear operator $T : V \to V$ and if $f \in \mathbb{F}[x]$, then the matrix of the operator $f(T) : V \to V$ is $f(A)$ with respect to the same basis. It follows that the minimal polynomial of T is same as the minimal polynomial of a matrix of T.

Note that T belongs to $\mathrm{Hom}_{\mathbb{F}}(V, V)$, which is of dimension n^2. Thus, $I, T, T^2, \ldots, T^{n^2}$ is a linearly dependent set and there exist scalars $a_0, a_2, \ldots, a_{n^2}$ such that

$$a_0 I + a_1 T + a_2 T^2 + \ldots + a_{n^2} T^n = O. \tag{2.14}$$

So, an annihilator of T is

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n^2} x^{n^2}$$

and we must have $\deg m_T(x) \leq n^2$.

Theorem 2.20. *Let $T : V \to V$ with n the dimension of the space V. The characteristic polynomial of T and the minimal polynomial of T have the same roots, except (possibly) for the multiplicities.*

Proof. We claim that $m_T(c) = 0$ if and only if c is an eigenvalue. Let $m_T(c) = 0$. Thus, $m_T(c) = (x - c) \cdot q(x)$, with $q \in \mathbb{F}[x]$ and $\deg q < \deg m$. Also, $q(T)$ is *not* the zero operator. So, there exists a $u \in V$ (non-zero vector) such that $q(T)(u) = v \neq 0$. Then,

$$0 = m(T)(u) = (T - cI) \cdot q(T)(u) = (T - cI)v \tag{2.15}$$

which shows that v is an eigenvector of T with eigenvalue c. So all roots of $m_T(x)$ are roots of the characteristic polynomial.

Conversely, let c be an eigenvalue of T. Say, $Tv = cv$ for some $v \neq 0$. Thus, $m_T(T)(v) = m(c)(v)$. But $m_T(T) = 0$ must mean that $0 = m(c)(v)$, and $m(c) = 0$. So every root of the characteristic polynomial is a root of the minimal polynomial. ■

January 30th.

Proposition 2.21. *If λ is an eigenvalue of T, then $f(\lambda)$ is an eigenvalue of $f(T)$ for $f \in \mathbb{F}[x]$.*

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

Proposition 2.22. *Let $T : V \to V$ be a diagonizable operator. The minimal polynomial is the product of distinct linear factors, that is, if*

$$\chi_T(x) = \prod_{i=1}^{k} (x - \lambda_i)^{d_i}$$

where the λ_i's are the distinct eigenvalues, then

$$m_T(x) = \prod_{i=1}^{k} (x - \lambda_i).$$

Proof. As T is a diagonalizable operator, there exists a basis of V consisting of eigenvectors of T, say $\mathcal{B} = (v_1, v_2, \ldots, v_n)$. Note that $m_T(T)v_i = 0$ for all valid i. For each $v_i \in \mathcal{B}$, there exists a λ_i such that $(T - \lambda_i I)v_i = 0$, which tells us

$$m_T(T) = (T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_k I)v_i = 0. \tag{2.16}$$

Hence, $m_T(x)$ is an annihilator for T, and it is of minimal degree by the above theorem. ■

Theorem 2.23 (*Cayley-Hamilton theorem*)**.** *Let $T : V \to V$ be a linear operator on a finite dimensional vector space V. If $\chi_T(x)$ is the characteristic polynomial of T, then $\chi_T(T) = 0$, that is, the characteristic polynomial annihilates T. Hence, the minimal polynomial of T divides the characteristic polynomial.*

Proof. Let $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ be a basis of V, and let $A = (a_{ij})$ be the matrix of T with respect to the basis $\mathcal{B}$. We have

$$\begin{aligned} a_{1j}v_1 + a_{2j}v_2 + \ldots + a_{nj}v_{nj} &= Tv_j \\ \implies -a_{1j}v_1 - a_{2j}v_2 - \ldots + (T - a_{jj})v_j - a_{(j+1)(j)}v_{j+1} - \ldots - a_{nj}v_n &= 0. \end{aligned} \tag{2.17}$$

This sysmte of equations can be written as

$$B_{n\times n}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.18}$$

where

$$B = \begin{pmatrix} T - a_{11}I & -a_{21} & \cdots & -a_{n1} \\ -a_{12} & T - a_{22}I & \cdots & -a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & T - a_{nn}I \end{pmatrix}. \tag{2.19}$$

Therefore $\det B = \chi_T(T)$. It is enough to show that $\det B = 0$ as an operator, that is, to show $\det B(b_i) = 0$ for all $v_i \in \mathcal{B}$. Let $(\text{adj}B)_{ij} = c_{ij}$, and $(B)_{ij} = b_{ij}$. Note that

$$\sum_{k=1}^{n} c_{ik}b_{kj} = \begin{cases} \det B & \text{if } i = j, \\ 0 & \text{if otherwise.} \end{cases}$$

Now,

$$\begin{aligned} &\sum_{j=1}^{n} b_{kj}v_j = 0 \text{ for all } 1 \leq k \leq n \\ \implies &\sum_{j=1}^{n} b_{kj}v_j = 0. \end{aligned}$$

Summing over all rows,

$$\begin{aligned} &\sum_{k=1}^{n}\left(\sum_{j=1}^{n} c_{ik}b_{kj}v_j\right) = 0 \\ \implies &\sum_{j=1}^{n}\left(\sum_{k=1}^{n} c_{ik}b_{kj}\right)v_j = 0. \end{aligned} \tag{2.20}$$

The left hand side is zero except for when $i = j$, in which case it is $\det B$—

$$0 = \sum_{j=1}^{n}\left(\sum_{k=1}^{n} c_{ik}b_{kj}\right)v_j = (\det B)v_i \tag{2.21}$$

which implies that the operator $\det B$ is zero on all the basis vectors, and hence it is the zero vecotr. Thus, since $\chi_T(T) = \det B$, $\chi_T(T)$ is also the zero operator. ■

Appendices

Chapter A

Appendix

Extra content goes here.

Appendix

Index

algebraic multiplicity, 10
annihilator, 11

Cayley-Hamilton theorem, 13
characteristic polynomial, 9
cofactor, 4
coordinate vector, 7
cycle, 1

determinant, 2
diagonizable linear operator, 10
diagonizable matrix, 10

eigenspace, 10
eigenvalue, 8
eigenvector, 8
even permutation, 2

geometric multiplicity, 10

ideal, 11

k-cycle, 1

minimal polynomial, 12

odd permutation, 2

principal ideal, 11

similar matrices, 8
symmetric group, 1

transpositions, 1