Lecture 2 Linear Algebra

We say λ is an eigenvalue of an $n \times n$ matrix A if there exists a vector $x \neq 0$ such that $Ax = \lambda x$

Equivalently λ is an eigenvalue of A if $\det(A - \lambda) = 0$

__Notation:__ We should properly write $\det(A - \lambda I) = 0$, but we will always write $\det(A - \lambda) = 0$ where the identity matrix I is understood.

If the underlying field $\mathbb{F}$ is algebraically closed (like $\mathbb{C}$, but not $\mathbb{R}$), the polynomial

$$P_A(\lambda) \equiv \det(A - \lambda)$$

always has a root. Hence eigenvalues always exist in this case

__Convention__ In what follows, whenever we talk about eigenvalues, we will __always__ assume that the underlying field is $\mathbb{C}$. It follows that A must in fact have n eigenvalues $\{\lambda_i\}_{i=1}^n$,

(17.0) $$p_A(\lambda) = (-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

The λ_i's need not be distinct: the number of times that λ_i occurs in the above product $p_A(\lambda)$, is called the <u>algebraic multiplicity</u> of λ_i. The <u>geometric multiplicity</u> of λ_i is the dimension of the null-space of $A-\lambda_i$, i.e. $\dim N(A-\lambda_i) = \dim\{x : (A-\lambda_i)x = 0\}$.

Recall that if A and B are square matrices of the same size n, then

(17.1) $$\det(AB) = (\det A)(\det B)$$

It follows that $\det A$ is invariant under conjugation, i.e. if U is an invertible $n\times n$ matrix, then

(17.2) $$\det \tilde{A} = \det A$$

where $\tilde{A} = UAU^{-1}$. Indeed $\det\tilde{A} = (\det U)(\det A)(\det U^{-1})$

$= \det A$, as $UU^{-1} = I$ implies $(\det U)(\det U^{-1}) = 1$.

It is this fact that allows us to assign an unambiguous notion of a determinant to a linear map A from a vector space V to itself. Indeed if $\{w_i\}_{i=1}^n$ is a basis for V, then $A w_i = \sum_{j=1}^n a_{ji} w_j$ from some matrix (a_{ij}), we define $\det A = \det(a_{ij})$. This is a good definition, because if $\{w_i'\}_{i=1}^n$ is another basis for V, then $A w_i' = \sum_{j=1}^n a_{ji}' w_j'$ for some matrix $\{a_{ij}'\}$. But a simple calculation shows that

$$(a_{ij}') = U(a_{ij})U^{-1}$$

for some invertible matrix U and hence

$$\det(a_{ij}') = \det(a_{ij})$$

by (17.2). Thus $\det A$ is well-defined independent of basis. In particular we can define an eigenvalue of a linear map $A : V \to V$ as before by saying λ

is an eigenvalue of A if there exists $x \in V$, $x \neq 0$, such that $Ax = \lambda x$. Because of the invariance of the determinant under conjugation it follows that we can compute the eigenvalues of A by solving the equation $\det((a_{ij}) - \lambda) = 0$ where (a_{ij}) is the matrix associated with A in any basis.

There is another matrix function that is of considerable interest, namely, for an $n \times n$ matrix $A = (a_{ij})$,

(19.1) $$\operatorname{tr} A \equiv \sum_{i=1}^{n} a_{ii}$$

In place of (17.1) we have

(19.2) $$\operatorname{tr} AB = \operatorname{tr} BA$$

for $n \times n$ matrices $A = (a_{ij})$, $B = (b_{ij})$. Indeed

$$\operatorname{tr} AB = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ji}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ji} A_{ij} = \sum_{j=1}^{n} (BA)_{jj} = \operatorname{tr} BA .$$

It follows that $\operatorname{tr} A$ is invariant under conjugation.

Indeed if $\tilde{A} = UAU^{-1}$ for some invertible matrix U

Then

$$\operatorname{tr} \tilde{A} = \operatorname{tr} UAU^{-1} = \operatorname{tr} U^{-1}UA = \operatorname{tr} A .$$

Thus we can define the trace, $\operatorname{tr} A$ of a linear map $A : V \to V$ for any vector space. We just note as before that the RHS of (19.1) is independent of the choice of basis for V.

Both $\det A$ and $\operatorname{tr} A$ are simply expressed in terms of the eigenvalues $\{\lambda_i\}$ of A. Indeed, we have from (17.0)

(20.1) $$p_A(\lambda) = \det(A - \lambda) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

But

$$\det(A - \lambda) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & & a_{n\,n-1} & a_{nn} - \lambda \end{pmatrix}$$

$$= (-\lambda)^n + (a_{11} + \cdots + a_{nn})(-\lambda)^{n-1} + \cdots + \det(a_{ij})$$

Comparing with (20.1) we find

(21.1) $$\det A = \lambda_1 \lambda_2 \cdots \lambda_n = \prod_{i}^{n} \lambda_i$$

and

(21.2) $$\operatorname{tr} A = \lambda_1 + \cdots + \lambda_n = \sum_{i=1}^{n} \lambda_i$$

Both the determinant and the trace play key roles in matrix theory / linear algebra.

What is the relationship between the algebraic and geometric multiplicities?

Insert from p21+

Lemma 21.3

If λ^0 is an eigenvalue of an $n \times n$ matrix $A = (a_{ij})$, then

$$1 \leq \text{geometric multiplicity of } \lambda^0 = \text{alg. mult. of } \lambda^0 \leq n$$

Proof: The only part of the result that remains to be proved is that

$$m \equiv \text{geom. mult.} \leq \text{algeb. mult.} \equiv k$$

To show this let $u_1, \ldots, u_m$ be a basis for $N(A - \lambda^0)$.

Extend this basis to a full basis $u_1, u_2, \ldots, u_m, u_{m+1}, \ldots u_n$ of $\mathbb{C}^n$ (see Exercise 21+1.2). As noted in (21+1.1), the matrix $A = (a_{ij})$ induces a

Insert on p21 (21+.1) If $(a_{ij})_{1\le i,j\le n}$, $a_{ij} \in \mathbb{C}$ (or $\mathbb{R}$), then (a_{ij}) define a linear mapping $A: \mathbb{C}^n \to \mathbb{C}^n$ by

$$Ae_i = \sum_{j=1}^{n} a_{ji} e_j , \quad 1 \le i \le n,$$

where $e_i = (0 \cdots 0 1 0 \cdots 0)^T$, $1 \le i \le n$, is the standard basis in $\mathbb{C}^n$ (or $\mathbb{R}^n$). Then if $x = (x_1, \dots, x_n)^T = \sum_{i=1}^{n} x_i e_i \in \mathbb{C}^n$, then

$$Ax = \sum_{i=1}^{n} x_i Ae_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ji} e_j$$

$$= \sum_{j=1}^{n} \Big(\sum_{i=1}^{n} a_{ji} x_i \Big) e_j$$

Thus the action of the induced mapping A is just standard matrix multiplication $x \to \sum_{i=1}^{n} a_{ji} x_i$, $j = 1, \dots, n$.

(21+1.2) Exercise: Show that if $u_1, \dots, u_m$ is an independent set of vectors in an n dimensional space V, then $u_1, \dots, u_m$ can be extended to a basis $u_1, \dots, u_m, u_{m+1}, \dots, u_n$ for V

(21+1.3) Exercise Show that if $u_1, \dots, u_m$ is an orthonormal set in $(V, (\cdot,\cdot))$, i.e. $(u_i, u_j) = \delta_{ij}$, $1 \le i \le j \le m$, then $u_1, \dots, u_m$ can be extended to an orthonormal basis $u_1, u_2, \dots, u_m, u_{m+1}, \dots, u_n$ for V.

linear map which we also denote by A, from $\mathbb{C}^n \to \mathbb{C}^n$.

Let $\tilde{A} = (\tilde{a}_{ij})$ be the matrix for A in the basis $u_1, \dots, u_n$, i.e. $A u_j = \sum_{i=1}^n \tilde{a}_{ij} u_i$.

If $u_j = \sum_{i=1}^n u_{ij} e_i$, then $U = (u_{ij})$ is invertible and

$\tilde{A} = U^{-1} A U$ (check this). Clearly $\det(\tilde{A} - \lambda) = \det(A - \lambda)$.

Now as $A u_\ell = \lambda^0 u_\ell$, $\ell = 1, \dots, m$, $\tilde{A}$ has the form

$$\begin{array}{c}
 \begin{array}{cc} \overset{m}{\longleftrightarrow} & \qquad\qquad \end{array} \\
 \begin{array}{c} \uparrow \\ m \\ \downarrow \\ \\ \end{array}
 \left(\begin{array}{ccc|c}
 \lambda^0 & & 0 & \\
 & \ddots & & X \\
 0 & & \lambda^0 & \\
 \hline
 & 0 & & X
 \end{array}\right) \\
 \begin{array}{cc} \qquad\qquad & \overset{n-m}{\longleftrightarrow} \end{array}
\end{array}$$

It follows that $\det(\tilde{A} - \lambda) = (\lambda^0 - \lambda)^m q(\lambda)$ where $q(\lambda)$

is a polynomial of order $n-m$. But $\det(A - \lambda) = (\lambda^0 - \lambda)^k s(\lambda)$

where $s(\lambda)$ is an $(n-k)^{th}$ order polynomial with $s(\lambda^0) \neq 0$.

Clearly we must have $m \leq k$, otherwise $s(\lambda^0) = 0$, which

is a contradiction. $\square$

Remark It is possible that geom. mult < alg. mult. For example, if

$$A = \begin{pmatrix} 0 & 1 & & 0 \\ 0 & 0 & 1 & \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{pmatrix}, \quad n \times n$$

then clearly $\det(A - \lambda I) = (-\lambda)^n$, so $\lambda = 0$ has alg. mult. $= n$. But if $Ax = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_2 \\ \vdots \\ x_n \\ 0 \end{pmatrix} = 0$,

then $x_i = 0$ for $i = 2, \ldots, n$. Thus $x = x_1 e_1$, and so

$\text{Nul } A = \langle e_1 \rangle = \text{span}\{e_1\}$, and so $1 =$ geom. mult. $<$ alg. mult. $= n$.

We will see, however, that for <u>normal</u> operators, i.e.

(23.1) $$AA^* = A^*A,$$

the two multiplicities are always equal. In particular this is true for Hermitian matrices $A = A^*$ and unitary matrices $AA^* = A^*A = I$.

We say an $n \times n$ matrix A is <u>diagonalizable</u> if there is a non-singular matrix $V = (v_{ij})$ and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$

such that

(24.1) $$A = V \Lambda V^{-1}$$

This means that in the basis

$$v_j = V e_j = \sum_{i=1}^{n} v_{ij} e_i \quad , \quad 1 \le j \le n .$$

A becomes diagonal, ie if $x = \sum_{j=1}^{n} x_j v_j$,

$$Ax = \sum_j x_j A v_j = \sum x_j A V e_j = \sum x_j V \Lambda e_j$$
$$= \sum x_j V \lambda_j e_j$$
$$= \sum \lambda_j x_j v_j$$

so

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \to \Lambda x = \begin{pmatrix} \lambda_1 x_1 \\ \vdots \\ \lambda_n x_n \end{pmatrix}$$

- The basic questions of linear algebra become trivial if A is a diagonal matrix, $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ say. For example, if we want to solve the equation $\Lambda x = b$, $\det \Lambda \neq$, then the solution is $x = \begin{pmatrix} b_1/\lambda_1 \\ \vdots \\ b_n/\lambda_n \end{pmatrix}$. Diagonalizable matrices are equally easy to handle: Just solve $Ax = b$ in the co-ordinates which diagonalize A and then return to the original co-ordinates. In a formula:

$$x = V \Lambda^{-1} (V^{-1} b)$$

(Note that if $b = \sum_{i=1}^{n} \tilde{b}_i v_i$, then $V^{-1} b = \sum_i \tilde{b}_i e_i = \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \end{pmatrix}$ which are the co-ordinates of b in the diagonalizing basis $v_1, \dots, v_n$.)

However, not every matrix is diagonalizable. For example if $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $A = V \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} V^{-1}$, then $\det(A - \lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) = \lambda^2$ and so $\lambda_1 = \lambda_2 = 0$. But then $A = 0$. Contradiction. It is of great and fundamental interest in linear algebra to determine classes of matrices that can be diagonalized.

There is the following interesting result.

Thm A is diagonalizable $\Leftrightarrow$ geom. mult. (λ_i) = alg. mult. (λ_i) for all eigenvalues λ_i

Proof: $\Rightarrow$ If $A = V\Lambda V^{-1}$, $\Lambda = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$

then $\det(A-\lambda) = \det(\Lambda - \lambda) = (\alpha_1 - \lambda)\cdots(\alpha_n - \lambda)$ so that

the α_i's are necessarily the eigenvalues of A. For any

$\alpha \in \{\alpha_1, \ldots, \alpha_n\}$, alg. mult $(\alpha) = \#\{i : \alpha_i = \alpha\}$. But

for any $j \in \{i : \alpha_i = \alpha\}$, we have from $AV = V\Lambda$

$$A v_j = \alpha v_j$$

where $v_j = Ve_j$ is the jth column of V. As V is invertible,

these v_j's are independent. It follows that

$$\text{alg. mult}(\alpha) = \#\{v_j = Ve_j : Av_j = \alpha v_j\} \leq \text{geom. mult}(\alpha)$$

But by Lemma 21.3, geom. mult. $(\alpha) \leq$ alg. mult (α), and hence

$$\text{geom. mult.}(\alpha) = \text{alg. mult.}(\alpha).$$

If geom. mult (λ) = alg. mult. (λ) for all eigenvalues, then

as the sum of the algebraic multiplicities of all the eigenvalues of

A is clearly n, it follows that

(26.1) $$\sum_{\text{all eig's } \lambda_i} \dim(N(A-\lambda_i)) = n$$

Arrange the eigenvalues in l distinct groups such that

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{k_1} \; ; \; \lambda_{k_1+1} = \cdots = \lambda_{k_2} \; ; \; \cdots \; ; \; \lambda_{k_{\ell-1}+1} = \cdots = \lambda_{k_\ell}$$

distinct

(clearly $k_\ell = n$) and let

$$v_1, \ldots, v_{k_1} \; ; \; v_{k_1+1}, \ldots, v_{k_2} \; ; \; \cdots \; ; \; v_{k_{\ell-1}+1}, \ldots, v_{k_\ell}$$

be bases for the associated nulspaces. We will show that $v_1, \ldots, v_n$ is a basis for $\mathbb{C}^n$. It is enough to show that the v_i's are independent. So suppose that for some $a_1, \ldots, a_n$, we have

$$\sum_{i=1}^{k_1} a_i v_i + \sum_{i=k_1+1}^{k_2} a_i v_i + \cdots = 0 . \tag{27.1}$$

Then after acting by A, we get

$$\lambda_{k_1} \sum_{i=1}^{k_1} a_i v_i + \lambda_{k_2} \sum_{i=k_1+1} a_i v_i + \cdots = 0 \tag{27.2}$$

and after multiplying (27.1) by λ_{k_1} and subtracting, we get

$$(\lambda_{k_1} - \lambda_{k_2}) \sum_{i=k_1+1}^{k_2} a_i v_i + (\lambda_{k_1} - \lambda_{k_3}) \sum_{i=k_2+1}^{k_3} a_i v_i + \cdots = 0 \tag{27.3}$$

~~As $(\lambda_{k_1} - \lambda_{k_2}) \neq 0$, Then by the appropriate induction assumption on ℓ, $0 = (\lambda_{k_1} - \lambda_{k_2}) a_{k_1+1} = \cdots = (\lambda_{k_1} - \lambda_{k_2}) a_{k_2} = (\lambda_{k_1} - \lambda_{k_3}) a_{k_2+1} = \cdots = (\lambda_{k_1} - \lambda_{k_3}) a_{k_3} = \cdots$ and as the λ_{k_i}'s are distinct, $a_i = 0$, $i > k_1$. But then $\sum_{i=1}^{k_1} a_i v_i = 0$, But as $v_1, \ldots, v_{k_1}$ is a basis for $N(A - \lambda_{k_1})$, $a_1 = \cdots = a_{k_1} = 0$. Thus the v_i's form a basis for $\mathbb{C}^n$. Set $V = (v_1, \ldots, v_n)$ and we find $A = V \Lambda V^{-1}$.~~

Then by induction on l, we get

$$(\lambda_{k_1}-\lambda_{k_2})a_{k_1+1} = \cdots = (\lambda_{k_1}-\lambda_{k_2})a_{k_2} = (\lambda_{k_1}-\lambda_{k_3})a_{k_2+1} = \cdots = (\lambda_{k_1}-\lambda_{k_3})a_{k_3} =$$

$$= \cdots = (\lambda_{k_1}-\lambda_{k_l})\cdot a_{k_l} = 0$$

However, as the λ_{k_i}'s are distinct, we conclude that $a_i = 0$ for $i > k_1$. But then from (27.1) we have $\sum_{i=1}^{k_1} a_i v_i = 0$, and as $v_1, \ldots, v_{k_1}$ are a basis for $\mathrm{Nul}(A-\lambda_{k_1})$, it follows that $a_1 = \cdots = a_{k_1} = 0$. Hence $v_1, \ldots, v_n$ are independent and so form a basis for $\mathbb{C}^n$. Set $V = (v_1 \cdots v_n)$ and we find $A = V \Lambda V^{-1}$. $\square$

Corollary 28.1 If all the eigenvalues of A are distinct, then A is diagonalizable.