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

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

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

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

$$P_A(\lambda) = \det(A - \lambda I)$$

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 $F$ is algebraically closed.
field is \( \mathbb{C} \). It follows that \( A \) must in fact have \( n \) eigenvalues \( \{ \lambda_1, \lambda_2, \ldots, \lambda_n \} \).

\[
P_A(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) .
\]

The \( \lambda \)'s need not be distinct: the number of times that \( \lambda_i \) occurs in the above product \( P_A(\lambda) \), is called the \underline{algebraic multiplicity} of \( \lambda_i \). The \underline{geometric multiplicity} of \( \lambda_i \) is the dimension of the null-space of \( A - \lambda_i I \), i.e., \( \dim \ker(A - \lambda_i I) = \dim \{ x : (A - \lambda_i I)x = 0 \} \).

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

\[
\det(AB) = (\det A)(\det B)
\]

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

\[
\det \tilde{A} = \det A
\]

where \( \tilde{A} = UAU^{-1} \). Indeed \( \det \tilde{A} = (\det U)(\det A)(\det U^{-1}) \).
\[ = \det A, \quad \text{so } uu^T = I \implies (\det u)(\det u^T) = 1. \]

It is the 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 \( \{w_1, \ldots, w_n\} \) is a basis for \( V \), then \( Aw_i = \sum_{j=1}^n a_{ij} w_j \) for some matrix \((a_{ij})\), we define \( \det A = \det (\alpha_{ij}) \). This is a good definition because if \( \{v_1, \ldots, v_n\} \) is another basis for \( V \), then \( Av_i = \sum_{j=1}^n a_{ij}' v_j \) 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...
In 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 $\det (a_{ii} - \lambda) = 0$, where $(a_{ii})$ 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 $A = (a_{ii})$,

$$\text{tr} \ A = \sum_{i=1}^{n} a_{ii} \quad (14.1)$$

In place of (17.1) we have

$$\text{tr} \ AB = \text{tr} \ BA \quad (14.2)$$

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

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

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} B_{ji} A_{ij} = \sum_{j=1}^{n} (BA)_{jj} = \text{tr} \ BA.$$
It follows that $\mathrm{tr} \ A$ is invariant under conjugation.

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

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

Thus we can define the trace, $\mathrm{tr} \ A$, 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 $\mathrm{tr} \ A$ are simply expressed in terms of the eigenvalues $\lambda_i$ of $A$. Indeed, we have from (19.0)

$$
(20.1) \quad p_A(\lambda) = \det (A - \lambda I) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)
$$

But

$$
\det (A - \lambda I) = \det \begin{pmatrix}
\lambda - a_{11} & a_{12} & \cdots & a_{1n} \\
\vdots & \ddots & \ddots & \vdots \\
\cdots & \cdots & \lambda - a_{n-1,n} & a_{nn} \\
0 & \cdots & \cdots & \lambda - a_{nn}
\end{pmatrix}
$$
\[
(-\lambda)^n + (a_{11} + \ldots + a_{nn})(-\lambda)^{n-1} + \ldots + \det(a_{ij})
\]

Comparing with (20.1) we find

\[
\det A = \lambda_1 \lambda_2 \ldots \lambda_n = \prod \lambda_i
\]

and

\[
\sum A = \lambda_1 + \ldots + \lambda_n = \sum \lambda_i
\]

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

What is the relationship between the algebraic and geometric multiplicities?

**Lemma 21.3**

If \( \lambda^0 \) 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

\[ m = \text{geom. mult.} \leq \text{alg. mult.} \leq k \]

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

Extend this basis to a full basis \( u_1, u_2, \ldots, u_m, u_{m+1}, \ldots, u_n \)

in \( \mathbb{C}^n \) (see Exercise 21.12). As noted in (21.1.1), the matrix \( A = (a_{ij}) \) induces a
A linear mapping \( A : \mathbb{C}^n \to \mathbb{C}^n \) by

\[
A e_i = \sum_{j=1}^{n} a_{ij} e_j , \quad i = 1, \ldots, n,
\]

where \( e_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0)^T \), \( i = 1, \ldots, n \), \( e_i \) in the standard basis \( e_1 \) in \( \mathbb{C}^n \) (or \( \mathbb{R}^n \)). Then if \( x = (x_1, \ldots, x_n)^T = \sum_{i=1}^{n} x_i e_i \in \mathbb{C}^n \),

\[
A x = \sum_{i=1}^{n} x_i A e_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} e_j = \sum_{j=1}^{n} \left( \sum_{i=1}^{n} a_{ij} x_i \right) e_j .
\]

Thus the action of the induced mapping \( A \) is just

Standard matrix multiplication \( x \to \sum_{i=1}^{n} a_{ij} x_i , \quad i = 1, \ldots, n \).

(214.1.2) Exercise: Show that if \( u_1, \ldots, u_m \) is an independent set of vectors in an \( n \)-dimensional space \( V \), then

\( u_1, \ldots, u_m \) can be extended to a basis \( u_1, \ldots, u_m, u_{m+1}, \ldots, u_n \) for \( V \).

(214.2) Exercise: Show that if \( u_1, \ldots, u_m \) is an orthonormal set in \( (V, (\cdot, \cdot)) \), i.e., \( (u_i, u_j) = \delta_{ij} \), \( 1 \leq i, j \leq m \), then \( u_1, \ldots, u_m \) can be extended to an orthonormal basis \( u_1, u_2, \ldots, u_m, u_{m+1}, \ldots, u_n \) 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 $u_1, \ldots, u_n$, i.e., $A u_i = \sum_j \tilde{a}_{ij} u_j$.

If $u_i = \sum_k e_k e_k^T$, then $U = (u_i e_j)$ is invertible and

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

Now as $A u_x = \lambda^x u_x$, $x = 1, \ldots, m$, $\tilde{A}$ has the form

\[
\begin{pmatrix}
\lambda^0 & \cdots & 0 \\
0 & \ddots & \vdots \\
0 & \cdots & \lambda^m
\end{pmatrix}
\begin{pmatrix}
x^0 \\
\vdots \\
x^m
\end{pmatrix}
\]

It follows that $\det(\tilde{A} - \lambda) = (\lambda^0 - \lambda)^m q(\lambda)$ where $q(\lambda)$ is a polynomial of order $n-m$, but $\det(A - \lambda) = (\lambda^0 - \lambda)^k s(\lambda)$ whose $s(\lambda)$ is an $(n-k)$th order polynomial with $s(\lambda^0) = 0$.

Clearly we must have $m \leq k$, otherwise $s(\lambda^0) = 0$, which is a contradiction. □

Remark: It is possible that geom. mult $< \text{alg. mult.}$, for example, if...
\[
A = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{pmatrix}, \quad n \times n
\]

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

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 no \( 1 \in \text{geom. mult.} \).

alg. mult. = n.

We will see, however, that for normal operators, i.e.

\[(23.1) \quad AA^* = A^*A,\]

the two multiplicities are always equal. In particular, this is true for Hermitian matrices \( A = A^* \) and unitary matrices \( A A^* = A^* A = I \).

We say an \( n \times n \) matrix \( A \) is \underline{diagonalizable} if there is a non-singular matrix \( V = (v_i) \) and a diagonal
Matrix $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}$

such that

$$A = V \Lambda V^{-1}$$

This means that in the basis

$$v_i = Ve_i = \sum_{i=1}^{n} \lambda_i e_i$$

A becomes diagonal, i.e.,

$$Ax = \sum_i x_i A e_i = \sum_i x_i \lambda_i e_i = \sum_i V \lambda_i e_i$$

Then,

$$x \Lambda^{-1} A x = \begin{pmatrix} \lambda_1 x_1 \\ \vdots \\ \lambda_n x_n \end{pmatrix}$$

The basic questions of linear algebra become trivial if $A$ is a diagonal matrix, $\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}$ say. For example, if we want to solve the equation $\Lambda x = b$, det $A$.

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 note \( Ax = b \)
in \( \mathbb{R}^n \) co-ordinates which diagonalize \( A \) and then return to the original co-ordinates. In a formula:

\[
x = V A^{-1} (V^T b)
\]

(Note that if \( b = \sum \delta_i \cdot v_i \), then \( V^T b = \sum \delta_i \cdot e_i \).

which are the co-ordinates of \( b \) in the diagonalizing basis \( v_1, \ldots, v_n \).

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 \( \det (A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) = \lambda^2 \) and no \( \lambda_1, \lambda_2 \) exist.

But then \( A \neq 0 \). Contradiction. It is of great fundamental interest in linear algebra to determine classes of matrices that can be diagonalized.

Here is the following interesting result.
Thm. A is diagonalizable \iff \text{geom. mult.} (\lambda_i) = \text{alg. mult.} (\lambda_i) \text{ for all eigenvalues } \lambda_i.

Proof: \implies \text{If } A = V \Lambda V^{-1}, \Lambda = \text{diag} (\lambda_1, \ldots, \lambda_n)

Then \det(A - \lambda I) = \det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda) = 0 \\text{ for each } \lambda_i.

As \lambda_i's are necessarily the eigenvalues of A, for any \lambda \in \{\lambda_1, \ldots, \lambda_n\}, \text{ alg. mult.} (\lambda) = \# \{i : \lambda_i = \lambda\}.

But for any \lambda \in \{\lambda_1, \ldots, \lambda_n\}, we have from \text{AV = VA}

\[A v_j = \lambda_i v_j\]

where \(v_j = V e_j\) is the j-th column of V. As V is invertible,

These \(v_j\)'s are independent. It follows that

\text{alg. mult.} (\lambda) = \# \{v_j = V e_j : A v_j = \lambda v_j\} \leq \text{geom. mult.} (\lambda).

But by Lemma 21.3, \text{geom. mult.} (\lambda) = \text{alg. mult.} (\lambda), and hence

\text{geom. mult.} (\lambda) = \text{alg. mult.} (\lambda) \text{ for all eigenvalues, then}

as the sum of the algebraic multiplicities of all the eigenvalues of A is clearly n, it follows that

\text{all eigenvalues } \chi \in \text{dim}(N(A - \lambda I)) = n

Arrange all eigenvalues in 2-dimension groups such that:

\[\sum \chi_i \in \text{dim}(N(A - \lambda I)) = n\]
\[ \lambda_1 = \lambda_2 = \ldots = \lambda_{k_1} \quad \lambda_{k_1+1} = \ldots = \lambda_{k_2} \quad \ldots \quad \lambda_{k_{k_1}} = \ldots = \lambda_{k_{k_2}} \]

(\text{clearly } k_2 = n) \text{ and let }

\[ v_1, \ldots, v_{k_1}, v_{k_1+1}, \ldots, v_{k_2}, \ldots, v_{k_{k_1}}, \ldots, v_{k_{k_2}} \]

be bases for the associated nullspaces. We will show that

\[ v_1, \ldots, v_n \text{ is a basis for } \mathbb{C}^n. \]  

It is enough to show that the \( v_i \)'s are independent. So suppose that

for some \( a_1, \ldots, a_n \), we have

\[
(27.1) \quad \sum_{i=1}^{k_1} a_i v_i + \sum_{i=k_1+1}^{k_2} a_i v_i + \ldots = 0.
\]

Then after acting by \( A \), we get

\[
(27.2) \quad \lambda_{k_1} \sum_{i=1}^{k_1} a_i v_i + \lambda_{k_2} \sum_{i=k_1+1}^{k_2} a_i v_i + \ldots = 0
\]

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

\[
(27.3) \quad (\lambda_{k_1} - \lambda_{k_2}) \sum_{i=k_1+1}^{k_2} a_i v_i + \ldots = 0.
\]

As \( (\lambda_{k_1} - \lambda_{k_2}) 
eq 0 \) then by the appropriate induction assumption

and as the \( \lambda_{k_1} \)'s are distinct, \( a_i = 0 \) for all \( i \). But then \( \sum_{i=1}^{k_1} a_i v_i = 0 \). Thus the \( v_i \)'s form a basis for \( \mathbb{C}^n \). Set \( V = \langle v_1, \ldots, v_n \rangle \) and we find \( A = UV \).
Then by induction on \( k \), we get

\[
(\lambda_{k_1} - \lambda_{k_2}) a_{k_1} = \cdots = (\lambda_{k_1} - \lambda_{k_k}) a_{k_1} = (\lambda_{k_1} - \lambda_{k_2}) a_{k_1+1} = \cdots = (\lambda_{k_1} - \lambda_{k_k}) a_{k_2} = \cdots = (\lambda_{k_1} - \lambda_{k_k}) a_{k_e} = 0
\]

However, as the \( \lambda_{k_i} \)'s are distinct, we conclude that \( a_{i} = 0 \) for \( i > k_1 \). But then from \((27.1)\) we have \( \sum_{i=1}^{k_1} a_i v_i = 0 \) if

and as \( v_1, \ldots, v_{k_1} \) are a basis for \( \text{Nul}(A - \lambda I) \), it follows that \( a_1 = \cdots = a_{k_1} = 0 \). Hence \( v_1, \ldots, v_{k_1} \) are independent and so form a basis for \( \mathbb{C}^{n} \). Set \( V = (v_1 \ldots v_{k_1}) \) and we find \( A = U \Lambda U^{-1} \). \( \square \)

**Corollary 28.1.** If all the eigenvalues of \( A \) are distinct, then \( A \) is diagonalizable.