EIGENVALUES AND EIGENVECTORS OF $n \times n$ MATRICES

With the formula for the determinant of a $n \times n$ matrix, we can extend our discussion on the eigenvalues and eigenvectors of a matrix from the $2 \times 2$ case to bigger matrices. To start we remind ourselves that an eigenvalue of $\lambda$ of $A$ satisfies the condition that $\det(A - \lambda I) = 0$, that is this new matrix is non-invertible.

**Proposition 0.1.** The eigenvalues of a square matrix $A$ are precisely the solutions $\lambda$ of the equation

$$\det(A - \lambda I) = 0$$

When we expand the determinant of the matrix $A - \lambda I$ we find a polynomial in $\lambda$, called the characteristic polynomial of $A$. The equation $\det(A - \lambda I) = 0$ is called the characteristic equation of $A$. Relating these facts to the $2 \times 2$ case, if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the characteristic polynomial is

$$\det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc)$$

If $A$ is a $n \times n$ matrix, the characteristic polynomial will be of degree $n$. From high-school algebra, the Fundamental theorem of algebra ensures that a polynomial of degree $n$ with real or complex coefficients has at most $n$ distinct roots, and so the characteristic polynomial of an $n \times n$ matrix with real or complex entries has at most $n$ distinct eigenvalues.

To summarize the procedure for finding eigenvalue and the corresponding eigenvectors for a matrix,

**Proposition 0.2.** Let $A$ be a $n \times n$ matrix

(1) Compute the characteristic polynomial $\det(A - \lambda I)$ of $A$.

(2) Find the eigenvalues of $A$ by solving the characteristic equation $\det(A - \lambda I) = 0$ for $\lambda$.

(3) For each eigenvalue $\lambda$, find the null space of the matrix $A - \lambda I$. This will be the eigenspace $E_\lambda$, that is the subspace of all non-zero vectors which are eigenvectors of $A$ corresponding to $\lambda$.

(4) Find a basis for each eigenspace.

**Example 0.3.** Q: Find the eigenvalues and corresponding eigenspaces for

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$$
A: To start we compute the characteristic polynomial

\[
det(A - \lambda I) = \begin{vmatrix} -\lambda & 0 \\ 0 & -\lambda & 1 \\ 2 & -5 & 4 - \lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & 1 \\ -5 & 4 - \lambda \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 2 & 4 - \lambda \end{vmatrix} = -\lambda(\lambda^2 - 4\lambda + 5) + 2 = -\lambda^3 + 4\lambda^2 - 5\lambda + 2
\]

the eigenvalues will be the roots of this polynomial equated to zero, that is, the roots of the characteristic equation \(\det(A - \lambda I) = 0\). This factors as \(-(\lambda - 1)^2(\lambda - 2)\), equating it to zero we find that \(\lambda = -1\) and \(\lambda = 2\) eigenvalues as these are both zeroes of the polynomial. Notice that \(\lambda = 1\) has multiplicity 2, while \(\lambda = 2\) is a simpler root; we index these values as \(\lambda_1 = \lambda_2 = 1\) and \(\lambda_3 = 2\). To find the eigenvectors corresponding to \(\lambda_1 = \lambda_2 = 1\) we find the null space of

\[
A - \lambda I = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & -5 & 3 \end{bmatrix}
\]

Combining this to form the augmented matrix \([A - I|0]\), row reduction yields

\[
\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}
\]

Hence if \(x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\) belongs to the eigenspace \(E_1\) it belongs to the null space of \(A - I\), with \(x_1 = x_3\) and \(x_2 = x_3\). Setting the free variable \(x_3 = t\) we find that the eigenspace is

\[
E_1 = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} = \text{span} \left( \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)
\]

Repeating this procedure to find the eigenvectors corresponding to \(\lambda_3 = 2\) we form the augmented matrix \([A - 2I|0]\) and row reduce:

\[
\begin{bmatrix} -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 2 & -5 & 2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1/2 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\]

Again supposing that \(x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\) is in the eigenspace \(E_2\), \(x_1 = \frac{1}{2}x_3\) and \(x_2 = \frac{1}{2}x_3\). Setting \(x_3 = t\) we find that

\[
E_2 = \left\{ t \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\} = \text{span} \left( \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right).
\]
In the previous example we worked with a $3 \times 3$ matrix, which had only two distinct eigenvalues. If we count multiplicities, the matrix $A$ has exactly three eigenvalues $\lambda = 1, 1, \text{ and } 2$. We define the **algebraic multiplicity** of an eigenvalue to be its multiplicity as a root of the characteristic equation of $A$. The next thing to note is that each eigenvector of $A$ has an eigenspace with a basis of one vector, so that $\dim E_1 = \dim E_2 = 1$. We define the **geometric multiplicity** of an eigenvalue $\lambda$ to be $\dim E_\lambda$, the dimension of its corresponding eigenspace. The connection between these two ideas of multiplicity will be important.

**Example 0.4.** Q: Find the eigenvalues and the corresponding eigenspaces of

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$$

A: In this case the characteristic equation is

$$\begin{vmatrix} -1 - \lambda & 0 & 1 \\ 3 & -\lambda & -3 \\ 1 & 0 & -1 - \lambda \end{vmatrix} = -\lambda \begin{vmatrix} -1 - \lambda & 1 \\ 1 & -1 - \lambda \end{vmatrix} = 0$$

$$= -\lambda(\lambda^2 + 2\lambda) = -\lambda^2(\lambda + 2)$$

thus the eigenvalues are $\lambda_1 = \lambda_2 = 0$ and $\lambda_3 = -2$. The eigenvalue 0 has algebraic multiplicity 2 and the eigenvalue -2 has algebraic multiplicity 1. For $\lambda_1 = \lambda_2 = 0$ row reduction of the matrix $[A - 0I|0]$ yields

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

from which it follows that any $x$ in $E_0$ satisfies $x_1 = x_3$, thus setting $x_2 = s$ and $x_3 = t$ we find

$$E_0 = s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \text{span} \left( \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right).$$

For $\lambda_3 = -2$ we find row reduction of $[A - (-2)I|0]$ yields

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

implying that any vector $x$ in the eigenspace $E_{-2}$ has $x_1 = x_3$ and $x_2 = 3x_3$, calling $x_3 = s$ we find:

$$E_{-2} = t \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} = \text{span} \left( \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right).$$

It follows that $\lambda_1 = \lambda_2 = 0$ has geometric multiplicity 2 and $\lambda_3 = -2$ has geometric multiplicity 1.

When we work with upper or lower triangular matrices, the eigenvalues of a matrix are very easy to find, if $A$ is triangular, then so is $A - \lambda I$ will be as well.
Thus the determinant of $A - \lambda I$ will be the product of the main diagonal entries. Thus

$$det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = 0$$

from which we immediately find that $\lambda_1 = a_{11}, \cdots, \lambda_n = a_{nn}$ are all eigenvalues.

**Theorem 0.5.** The eigenvalues of a triangular matrix are the entries on its main diagonal.

**Example 0.6.** Q: Find the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 5 & 7 & 4 & -2 \end{bmatrix}$$

A: These are just the entries along the diagonal, hence $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 3$ and $\lambda_4 = -2$

Eigenvalues encode important information about the behaviour of a matrix. Once we know the eigenvalues of a matrix we can determine many helpful facts about the matrix without doing any more work.

**Theorem 0.7.** A square matrix $A$ is invertible if and only if 0 is not an eigenvalue of $A$.

**Proof.** Let $A$ be a square matrix, we now know that a matrix is invertible if and only if its determinant is nonzero, i.e. $\det A \neq 0$. This condition is equivalent to $\det(A - 0I) = 0 \neq 0$ implying that 0 is not a root of the characteristic equation of $A$, and hence cannot be an eigenvalue. \[ \square \]

with this theorem we may extend the Fundamental Theorem of Invertible Matrices

**Theorem 0.8.** The Fundamental Theorem of Invertible Matrices: Ver. 3 Let $A$ be an $n \times n$ matrix. the following statements are equivalent (i.e. all true or all false):

1. $A$ is invertible
2. $Ax = b$ has a unique solution for every $b$ in $\mathbb{R}^n$
3. $Ax = 0$ has only the trivial solution.
4. The reduced row echelon form of $A$ is the $n \times n$ identity matrix
5. $A$ is a product of elementary matrices.
6. $\text{rank}(A) = n$
7. $\text{nullity}(A) = 0$
8. The column vectors of $A$ are linearly independent.
9. The column vectors of $A$ span $\mathbb{R}^n$
10. The column vectors of $A$ form a basis for $\mathbb{R}^n$
11. The row vectors of $A$ are linearly independent
12. The row vectors of $A$ span $\mathbb{R}^n$
13. The row vectors of $A$ form a basis for $\mathbb{R}^n$
14. $\det A \neq 0$
15. 0 is not an eigenvalue of $A$.

Invoking eigenvalues, there are nice formulas for the powers and inverses of a matrix
Theorem 0.9. Let $A$ be a square matrix with eigenvalue $\lambda$ and corresponding eigenvector $x$.

1. For any positive integer $n$, $\lambda^n$ is an eigenvalue of $A^n$ with corresponding eigenvector $x$.
2. If $A$ is invertible, then $1/\lambda$ is an eigenvalue of $A^{-1}$ with corresponding eigenvector $x$.
3. If $A$ is invertible, then for any integer $n$, $\lambda^n$ is an eigenvalue of $A^n$ with corresponding eigenvector $x$.

Proof. We proceed by induction on $n$; for the base-case $n = 1$ the result is what has been given, for the inductive-assumption we assume the result is true for $n = k$, $A^k x = \lambda^k x$. To prove this for arbitrary $n$, we must show this holds for $n = k + 1$.

As the identity $A^{k+1} x = A(A^k x) = A(\lambda^k x)$ - using the inductive assumption, we find that

$$A(\lambda^k x) = \lambda^k (Ax) = \lambda^k (k x) = \lambda^{k+1} x$$

Thus $A^{k+1} x = \lambda^{k+1} x$ holds for arbitrary $k$, proving that this is indeed true for any positive integer. The remaining two properties may be proven in a similar manner. 

Example 0.10. Q: Compute the matrix product

$$\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}^{10} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

A: Let $A$ be the matrix and $x$ be the vector; then we wish to compute $A^{10} x$.

The eigenvalues of $A$ are $\lambda_1 = -1$ and $\lambda_2 = 2$ with eigenvectors $v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, implying the following identities

$$Av_1 = -v_1, \quad Av_2 = 2v_2$$

As $v_1$ and $v_2$ are linearly independent they form a basis for $\mathbb{R}^2$ and so we may write $x$ as a linear combination of the two eigenvectors, $x = 3v_1 + 2v_2$. Thus applying the previous theorem we find

$$A^{10} x = A^{10} (3v_1 + 2v_2) = 3(A^{10} v_1) + 2(A^{10} v_2) = 3(-1)^{10} v_1 + 2(2)^{10} v_2$$

Expanding this out we find

$$3(-1)^{10} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 2(2^{10}) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2051 \\ 4093 \end{bmatrix}$$

Theorem 0.11. Suppose the $n \times n$ matrix $A$ has eigenvectors $v_1, v_2, ..., v_m$ with corresponding eigenvalues $\lambda_1, \lambda_2, ..., \lambda_m$. If $x$ is a vector in $\mathbb{R}^n$ that can be expressed as a linear combination of these eigenvectors. That is

$$x = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m$$

then for any integer $k$,

$$A^k x = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_m \lambda_m^k v_m$$
This may not always hold, there is no guarantee that such a linear combination is possible. The best possible situation would be if there were a basis of $\mathbb{R}^n$ consisting of eigenvectors of $A$, however this many not always be the case. The next theorem states that eigenvectors corresponding to distinct eigenvalues are linear independent.

**Theorem 0.12.** Let $A$ be an $n \times n$ matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be distinct eigenvalues of $A$ with corresponding eigenvectors $v_1, v_2, \ldots, v_m$. Then $v_1, v_2, \ldots, v_m$ are linearly independent.

**Proof.** We will use an indirect contradiction proof: suppose $v_1, v_2, \ldots, v_m$ are linear dependent - we will show a contradiction arises.

If $v_1, v_2, \ldots, v_m$ are linearly dependent, one of these vectors must be expressible as a linear combination of the remaining vectors. Let $v_{k+1}$ be the first of the vectors $v_1$ that may be expressed in this way. So that $v_1, v_2, \ldots, v_k$ are linearly independent, but that there are scalars $c_1, c_2, \ldots, c_k$ such that

$$v_{k+1} = c_1 v_1 + \ldots + c_k v_k$$

Multiplying both sides of this equation by $A$ from the left and using the fact that $A v_i = \lambda_i v_i$ we find

$$\lambda_{k+1} v_{k+1} = A v_{k+1} = A(c_1 v_1 + \ldots + c_k v_k) = c_1 A v_1 + \ldots + c_k A v_k = c_1 \lambda_1 v_1 + \ldots + c_k \lambda_k v_k.$$ 

Alternatively multiplying this equation by $\lambda_{k+1}$ we find

$$\lambda_{k+1} v_{k+1} = c_1 \lambda_{k+1} v_1 + \ldots + c_k \lambda_{k+1} v_k.$$ 

Subtracting these two equations from each other we find

$$0 = c_1 (\lambda_1 - \lambda_{k+1}) v_1 + \ldots + c_k (\lambda_k - \lambda_{k+1}) v_k.$$ 

The linear independence of $v_1, \ldots, v_k$ implies

$$c_1 (\lambda_1 = \lambda_{k+1}) = \ldots = c_k (\lambda_k - \lambda_{k+1}) = 0$$

Since the eigenvalues of $A$ are all distinct, the terms in the brackets are all non-zero. Thus $c_1 = c_2 = \ldots = c_k = 0$, and

$$v_{k+1} = c_1 v_1 + \ldots + c_k v_k = 0$$

this is impossible as an eigenvector is always non-zero. This is a contradiction implying that our assumption that $v_1, \ldots, v_m$ are linearly dependent is false and so these $m$ eigenvectors must be linearly independent. $\square$

**Similarity and Diagonalization**

We’ve seen that triangular and diagonal matrices have a useful property: their eigenvalues are easily read off along the diagonal. If we could relate a given square matrix to a triangular or diagonal matrix with the same eigenvalues, this would be incredibly useful. Of course, one could use Gaussian elimination, however this process will change the column space of the matrix and the eigenvalues will be altered as well. In the last section of the course we introduce a different transformation of a matrix that will not change the eigenvalues.
Similar Matrices.

**Definition 0.13.** Let $A$ and $B$ be $n \times n$ matrices, we say that $A$ is similar to $B$ if there is an invertible $n \times n$ matrix $P$ such that $P^{-1}AP = B$. If $A$ is similar to $B$ we write $A \sim B$.

Notice that $P$ depends on $A$ and $B$ and it is not unique for a given pair of matrices; for example if $A = B = I_n$ then $P^{-1}I_nP = I_n$ is satisfied for any invertible matrix $P$. Given a particular instance of $P^{-1}AP = B$ we may multiply on the left to produce the identity $AP = PB$.

**Example 0.14.** Q: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$, show that $A \sim B$.

A: Consider the matrix products $\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & -1 \end{bmatrix}$

Then $AP = PB$ with $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

**Theorem 0.15.** Let $A, B,$ and $C$ be $n \times n$ matrices.

1. $A \sim A$.
2. If $A \sim B$ then $B \sim A$.
3. If $A \sim B$ and $B \sim C$ then $A \sim C$.

**Proof.**

1) This follows immediately since $I^{-1}AI = A$.

2) If $A \sim B$ then $P^{-1}AP = B$ for some invertible matrix $P$. Writing $Q = P^{-1}$ we find that this equation may be written as $Q^{-1}BQ = (P^{-1})^{-1}BP^{-1} = PBP^{-1} = A$, hence $B \sim A$.

3) Suppose $A = P^{-1}BP$ and $B = Q^{-1}CQ$ where $P$ and $Q$ are invertible matrices, then $A = P^{-1}Q^{-1}CQP$, denoting $N = QP$ we see that $N^{-1} = P^{-1}Q^{-1}$ and so $A = N^{-1}CN$ proving that $A \sim C$.

Any relation satisfying these three properties is called an equivalence relation, these appear in many areas of mathematics where certain objects are related under some equivalence relation - usually where they share similar properties. We will see an example of this with similar matrices.

**Theorem 0.16.** Let $A$ and $B$ be $n \times n$ matrices with $A \sim B$, then

1. $\det A = \det B$.
2. $A$ is invertible if and only if $B$ is invertible.
3. $A$ is invertible if and only if $B$ is invertible.
4. $A$ and $B$ have the same rank.
5. $A$ and $B$ have the same characteristic polynomial.
6. $A$ and $B$ have the same eigenvalues.

**Proof.** We prove 1) and 4), and leave the remaining properties as exercises. Recall that if $A \sim B$ then $P^{-1}AP = B$ for some invertible matrix $P$.

1) Taking determinants on both sides we find

$$\det B = \det(P^{-1}AP) = \det P^{-1} \det A \det P = \frac{1}{\det P} \det B \det A = \det A$$
2) The characteristic polynomial of $B$ is
\[
det(B - \lambda I) = \det(P^{-1}AP - \lambda I) = \det(P^{-1}AP - \lambda P^{-1}IP) = \det(P^{-1}AP - P^{-1}\lambda P) = \det(P^{-1}(A - \lambda I)P) = \det(A - \lambda I)\]
\[
\square
\]
This theorem is helpful for proving if two matrices are not similar, as there are matrices which have all properties 1-5 in common and yet are not similar. Consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ both have determinant 1 and rank 2, are invertible and have characteristic polynomial $(1 - \lambda)^2$ and eigenvalues $\lambda_1 = \lambda_2 = 1$ - but these two matrices are not similar since $P^{-1}AP = P^{-1}IP = I \neq B$ for any invertible matrix $P$.

Example 0.17. Consider the pairs of matrices:

- $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ are not similar since $\det A = -3$ but $\det B = 3$.
- $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}$ are not similar since the characteristic polynomial of $A$ is $\lambda^2 - 3\lambda - 4$ while $B$ has $\lambda^2 - 4$.

**Diagonalization.** The best we can hope for when we are given a matrix is when it is similar to a diagonal matrix. In fact there is a close relationship between when a matrix is diagonalizable and the eigenvalues and eigenvectors of a matrix.

**Definition 0.18.** A $n \times n$ matrix $A$ is **diagonalizable** if there is a diagonal matrix $D$ such that $A$ is similar to $D$, i.e. there is some $n \times n$ invertible matrix $P$ such that $P^{-1}AP = D$.

Example 0.19. The matrix $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ is diagonalizable since the matrix $P = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ produce the identity $P^{-1}AP = D$ or equivalently $AP = AD$.

This is wonderful, but we have no idea where $P$ and $D$ arose from. To answer this question we note that the diagonal entries 4 and -1 of $D$ are the eigenvalues of $A$ since they are roots of its characteristic polynomial so we have an idea where $D$ is coming from. How $P$ is found is a more interesting question, as in the case of $D$, the entries of $P$ are related to the eigenvectors of $A$.

**Theorem 0.20.** Let $A$ be an $n \times n$ matrix, then $A$ is diagonalizable if and only if $A$ has $n$ linearly independent eigenvectors.

To be precise, there exists an invertible matrix $P$ and a diagonal matrix $D$ such that $P^{-1}AP = D$ if and only if the columns of $P$ are $n$ linearly independent eigenvectors of $A$ and the diagonal entries of $D$ are the eigenvalues of $A$ corresponding to the eigenvectors in $P$ in the same order.

**Proof.** Suppose that $A$ is similar to the diagonal matrix $D$ via $AP = PD$, and let the columns of $P$ be $p_1, p_2, ..., p_n$ and let the diagonal entries of $D$ be $\lambda_1, \lambda_2, ..., \lambda_n$. 

Then
\[
A[p_1 \; p_2 \; \cdots \; p_n] = [p_1 \; p_2 \; \cdots \; p_n]
\]
\[
\begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{pmatrix}
\]
\[
A[p_1 \; Ap_2 \; \cdots \; Ap_n] = [\lambda_1 p_1 \; \lambda_2 p_2 \; \cdots \; \lambda_n p_n]
\].

Equating each column we find that for \(i = 1, 2, \ldots, n\),
\[Ap_i = \lambda_i p_i\]
proving that the column vectors of \(P\) are eigenvectors of \(A\) whose corresponding eigenvalues are the diagonal entries of \(D\) in the same order. As \(P\) is invertible its columns are linearly independent by the Fundamental Theorem of Invertible Matrices.

On the other hand, if \(A\) has \(n\) linearly independent eigenvectors \(p_1, p_2, \ldots, p_n\) with corresponding eigenvalues \(\lambda_1, \lambda_2, \ldots, \lambda_n\) respectively then
\[Ap_i = \lambda_i p_i, \quad i = 1, 2, \ldots, n.\]

This statement leads to the original matrix product \(AP = DP\) where \(P\) is the matrix formed by making the eigenvectors column vectors of the matrix and \(D\) the diagonal matrix with the eigenvalues as entries. As the columns of \(P\) are linearly independent the Fundamental Theorem of Invertible Matrices it will be invertible and so \(P^{-1}AP = D\) proving \(A\) is diagonalizable. \(\square\)

**Example 0.21.** Q: Determine whether a matrix \(P\) exists to diagonalize
\[
A = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
2 & -5 & 4
\end{bmatrix}
\].

A: We have seen that this matrix has eigenvalues \(\lambda_1 = \lambda_2 = 1\) and \(\lambda_3 = 2\) with the corresponding eigenspaces
\[E_1 = \text{Span} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad E_2 = \text{Span} \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.
\]

As all other eigenvectors are just multiples of one of these two basis vectors there cannot be three linearly independent eigenvectors. Thus it cannot be diagonalized.

**Example 0.22.** Q: Find a \(P\) that will diagonalize
\[
A = \begin{bmatrix}
-1 & 0 & 1 \\
3 & 0 & -3 \\
1 & 0 & -1
\end{bmatrix}
\].

A: Previously we had seen that this matrix has eigenvalues \(\lambda_1 = \lambda_2 = 0\) and \(\lambda_3 = -2\) with the basis for the eigenspaces:
\[E_0 = \text{span} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad E_{-2} = \text{span} \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}.
\]
It is easy to check that the three vectors are linearly independent, and so we form
\[ P = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & 1 \end{bmatrix} \]
this matrix will be invertible and furthermore
\[ P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} = D. \]

In the last example we checked to see if the three eigenvectors are linearly independent, but was this necessary? We knew that the first two basis eigenvectors in the eigenspace for 0 were linearly independent but how do we know the pairing of one basis vector from either eigenspace will be linearly independent? The next theorem resolves this issue.

**Theorem 0.23.** Let \( A \) be an \( n \times n \) matrix and let \( \lambda_1, \lambda_2, \ldots, \lambda_k \) be distinct eigenvalues of \( A \). If \( \mathcal{B}_i \) is a basis for the eigenspace of \( E_{\lambda_i} \), then \( \mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_k \) is linearly independent.

**Proof.** Let \( \mathcal{B} = \{v_{i1}, v_{i2}, \ldots, v_{in_i}\} \) for \( i = 1, \ldots, k \) we must show that
\[ \mathcal{B} = \{v_{11}, v_{12}, \ldots, v_{1n_1}, v_{21}v_{22}, \ldots, v_{2n_2}, \ldots, v_{k1}v_{k2}, \ldots, v_{kn_k}\} \]
is linearly independent. Suppose some non-trivial linear combination of these vectors is the zero vector
\[ (c_{11}v_{11} + \ldots + c_{1n_1}v_{1n_1}) + (c_{21}v_{21} + \ldots + c_{2n_1}v_{2n_1}) + \ldots + (c_{k1}v_{k1} + \ldots + c_{kn_1}v_{kn_1}) = 0 \]
Expressing the sums in brackets by \( x_1, x_2, \ldots, x_k \) we may write this as
\[ x_1 + x_2 + \ldots + x_k = 0. \]
Now each \( x_i \) belongs in \( E_{\lambda_i} \) and so either is an eigenvector corresponding to \( \lambda_i \) or it is the zero vector. As the eigenvalues \( \lambda_i \) are distinct, if any of the factors \( x_i \) is an eigenvector they are linearly independent. However, the above is a linear independence relationship and so this must be a contradiction; we conclude that \( \mathcal{B} \) is linearly independent. \( \square \)

There is one case where diagonalizability is automatic, the case where the matrix \( A \) has \( n \) eigenvalues

**Theorem 0.24.** If \( A \) is a \( n \times n \) matrix with \( n \) distinct eigenvalues, then \( A \) is diagonalizable.

**Proof.** Let \( v_1, v_2, \ldots, v_n \) be eigenvectors corresponding to the \( n \) distinct eigenvalues of \( A \). These vectors are linearly independent by Theorem (0.11) and so Theorem (0.20) \( A \) is diagonalizable. \( \square \)

**Example 0.25.** The matrix
\[ A = \begin{bmatrix} 2 & -3 & 7 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{bmatrix} \]
has eigenvalues \( \lambda_1 = 2, \lambda_2 = 5 \) and \( \lambda_3 = -1 \), as these eigenvalues are distinct for the \( 3 \times 3 \) matrix \( A \), \( A \) is diagonalizable by the last theorem.
As a final theorem we characterize diagonalizable matrices in terms of two notions of multiplicity: algebraic and geometric. We will give precise conditions under which a \( n \times n \) matrix can be diagonalized, even when it has fewer eigenvalues than the size of the square matrix. To do this we prove a helpful lemma first.

**Lemma 0.26.** If \( A \) is an \( n \times n \) matrix, then the geometric multiplicity of each eigenvalue is less than or equal to its algebraic multiplicity.

**Proof.** Suppose \( \lambda_1 \) is an eigenvalue of \( A \) with geometric multiplicity \( p \), so that \( \dim E_{\lambda_1} = p \). Supposing that this eigenspace has the basis \( B = \{ v_1, v_2, ..., v_p \} \). Let \( Q \) be any invertible \( n \times n \) matrix having \( v_1, v_2, ..., v_p \) as its first \( p \) columns

\[
Q = [v_1 \cdots v_p v_{p+1} \cdots v_n]
\]

or as a partitioned matrix \( Q = [U | V] \). We define

\[
Q^{-1} = \begin{bmatrix} C \\ D \end{bmatrix}
\]

where \( C \) is a \( p \times n \) matrix. As the columns of \( U \) are eigenvectors corresponding to \( \lambda_1 \), \( AU = \lambda_1 U \) and we also have

\[
\begin{bmatrix}
I_p & O \\
O & I_{n-p}
\end{bmatrix}
= I_n = Q^{-1}Q = \begin{bmatrix} C \\ D \end{bmatrix} [U | V] = \begin{bmatrix} CU \\ DU \end{bmatrix} \begin{bmatrix} CV \\ DV \end{bmatrix}
\]

from which we obtain that \( CU = I_p, CV = O, DU = O \) and \( DV = I_{n-p} \). Therefore

\[
Q^{-1}AQ = \begin{bmatrix} C \\ D \end{bmatrix} [U | V] = \begin{bmatrix} CAU & CAV \\ DAV & DAV \end{bmatrix} = \begin{bmatrix} \lambda_1 CU & \lambda_1 CV \\ \lambda_1 DU & \lambda_1 DV \end{bmatrix} = \begin{bmatrix} \lambda_1 I_p & CAV \\ O & DAV \end{bmatrix}
\]

It follows that

\[
\det(Q^{-1}AQ - \lambda I) = (\lambda_1 - \lambda)^p \det(DAV - \lambda I)
\]

but \( \det(Q^{-1}AQ - \lambda I) \) is the characteristic polynomial of \( Q^{-1}AQ \) which is the same as the characteristic polynomial for \( A \). Thus this implies that the algebraic multiplicity of \( \lambda_1 \) is at least \( p \), its geometric multiplicity. \( \square \)

**Theorem 0.27.** The Diagonalization Theorem Let \( A \) be an \( n \times n \) matrix whose distinct eigenvalues are \( \lambda_1, \lambda_2, ..., \lambda_n \), the following statements are equivalent:

1. \( A \) is diagonalizable.
2. The union \( \mathcal{B} \) of the bases of the eigenspace of \( A \) contains \( n \) vectors.
3. The algebraic multiplicity of each eigenvalue equals its geometric multiplicity.

**Proof.** To prove 1) \( \rightarrow \) 2) Suppose \( A \) is diagonalizable, then it has \( n \) linearly independent eigenvectors. If \( n_i \) of these eigenvectors correspond to the eigenvalue \( \lambda_i \) then \( \mathcal{B}_i \) contains at least \( n_i \) vectors. Thus \( \mathcal{B} \) contains at least \( n \) vectors, and this basis is linearly independent in \( \mathbb{R}^n \) it must contain exactly \( n \) vectors.

To show 2) \( \rightarrow \) 3), let the geometric multiplicity of \( \lambda_i \) be \( d_i = \dim E_{\lambda_i} \) and let the algebraic multiplicity of \( \lambda_i \) be \( m_i \). By the previous lemma \( d_i \leq m_i \) for \( i = 1, 2, ..., k \).

If we assume the second property holds then we also have

\[
n = d_1 + d_2 + ... + d_k \leq m_1 + m_2 + ... + m_k
\]
However \( m_1 + m_2 + \ldots + m_k = n \) since the sum of the algebraic multiplicities of the eigenvalues of \( A \) is the degree of the characteristic polynomial of \( A \) which is \( n \). Thus it follows that \( d_1 + d_2 + \ldots + d_k = m_1 + m_2 + \ldots + m_k = n \) which implies that 

\[
(m_1 - d_1) + (m_2 - d_2) + \ldots + (m_k - d_k) = 0
\]

Using the lemma again we know that \( m_i - d_i \geq 0 \) for \( i = 1, 2, \ldots, k \) from which we deduce that each term in the sum is zero and so \( m_i = d_i \) for \( i = 1, 2, \ldots, k \).

To show 3) \( \rightarrow \) 1) we note that if the algebraic multiplicity \( m_i \) and the geometric multiplicity \( d_i \) are equal for each eigenvalue \( \lambda_i \) of \( A \) then \( B \) has \( d_1 + d_2 + \ldots + d_k = m_1 + m_2 + \ldots + m_k = n \) vectors, which we now know are linearly independent. Thus there are \( n \) linearly independent eigenvectors of \( A \) and \( A \) is diagonalizable. \( \square \)

\textbf{Example 0.28.}  

\( \bullet \) The matrix \( A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & -5 & 4 \end{bmatrix} \) has two distinct eigenvalues, \( \lambda_1 = \lambda_2 = 1 \) and \( \lambda_3 = 2 \). Since the algebraic multiplicity of the eigenvalue 1 is 2 but its geometric multiplicity is 1 \( A \) is not diagonalizable by the Diagonalization Theorem.

\( \bullet \) The matrix \( A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix} \) has two distinct eigenvalues \( \lambda_1 = \lambda_2 = 0 \) and \( \lambda_3 = -2 \). The eigenvalue 0 has algebraic and geometric multiplicity 2 and the eigenvalue -2 has algebraic and geometric multiplicity 1. By the diagonalization theorem this matrix is diagonalizable.

We conclude this section with a helpful application of diagonalizable matrices

\textbf{Example 0.29.} Q: Compute \( A^{10} \) if

\[
A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}.
\]

A: We have seen that this matrix has eigenvalues \( \lambda_1 = -1 \) and \( \lambda_2 = 2 \) with corresponding eigenvectors \( v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \) and \( v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \). It follows that \( A \) is diagonalizable and \( P^{-1}AP = D \) where

\[
P = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \text{ and } D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}.
\]

Solving for \( A \) we have \( A = PDP^{-1} \), the powers of \( A \) are now easily expressed since

\[
A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1})DP^{-1} = PD(P^{-1}P)DP^{-1} = PDIDP^{-1}PD^2P^{-1}
\]

and in general \( A^n = PD^nP^{-1} \) for any \( n \geq 1 \), which is true for any diagonalizable matrix. Computing \( D^n \) we find

\[
D^n = \begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix}
\]

we have that

\[
A^n = PD^nP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2(-1)^n+2^n}{3} & \frac{(-1)^n+2^n}{3} \\ \frac{(-1)^n+2^n}{3} & \frac{2(-1)^n+2^n}{3} \end{bmatrix}.
\]
Choosing $n = 10$ we find that

$$A^{10} = \begin{bmatrix}
\frac{2(-1)^{10} + 2^{10}}{2(-1)^{11} + 2^{11}} & \frac{(-1)^{11} + 2^{10}}{2(-1)^{11} + 2^{11}} \\
\frac{3}{3} & \frac{3}{3} \\
\frac{682}{683} & \frac{682}{683}
\end{bmatrix} = \begin{bmatrix} 342 & 341 \\
682 & 683
\end{bmatrix}. $$

References