Note: Mathematical Induction uses the previous case to imply the next one, so if the first case holds, then they all must hold. Consider the following proof to demonstrate. We will show that for all positive integers $n$,

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$ 

First, we show it holds for $n = 1$. For $n = 1$, the above statement becomes $1 + 1 = 2$, which is true. Now, we will show that the previous result, for $n$, implies the next one, for $n + 1$. To do this, we assume the result holds for $n$. That is, we assume

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

This is called the Induction Hypothesis (IH). Next, we prove the next case for $(n+1)$ as follows:

$$1 + 2 + \cdots + n + (n+1) = (1 + 2 + \cdots + n) + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n^2 + n}{2} + \frac{2n+2}{2} = \frac{n^2 + 3n + 2}{2} = \frac{(n+1)(n+2)}{2}.$$ 

Notice that we used the induction hypothesis in the first step. The rest is just algebra. Also, we recognize $1 + 2 + \cdots + n + (n+1) = \frac{(n+1)(n+2)}{2}$ as the statement with $n$ replaced by $n + 1$. So, because it holds for $n = 1$, and each result implies the next one, it is true for all positive integers $n$. This is how induction works.

FOR PROBLEMS 1a.) AND 4 c.), I RECOMMEND USING INDUCTION, BUT IT IS NOT REQUIRED.

1. Suppose that $(\lambda, \vec{v})$ is an eigenpair of matrix $A$.

(a) Show $(\lambda^k, \vec{v})$ is an eigenpair of $A^k$, for all positive integers $k$.

(b) Assume $A$ is invertible. Why does this mean $\lambda \neq 0$? Then, show $(\lambda^{-1}, \vec{v})$ is an eigenpair of $A^{-1}$.

(a) We are given that $A\vec{v} = \lambda \vec{v}$. We proceed by induction on $k$. First for $k = 1$, we already know $A\vec{v} = \lambda \vec{v}$, which proves the base case. Now, we assume that $(\lambda^k, \vec{v})$ is an eigenpair of $A^k$, or $A^k \vec{v} = \lambda^k \vec{v}$. Then,

$$A^{k+1} \vec{v} = A(A^k \vec{v}) = A(\lambda^k \vec{v}) = \lambda^k (A \vec{v}) = \lambda^k (\lambda \vec{v}) = \lambda^{k+1} \vec{v},$$

so $(\lambda^{k+1}, \vec{v})$ is an eigenpair of $A^{k+1}$. This completes our induction.

(b) Suppose to the contrary that $\lambda$ is an eigenvalue of $A$. Then, $A\vec{v} = 0\vec{v} = \vec{0}$ for some $\vec{v} \neq 0$, which means $(A) \neq \{\vec{0}\}$, so $A$ is not invertible, a contradiction. Thus, $0$ cannot be an eigenvalue of $A$, so $\lambda \neq 0$. Then,

$$A\vec{v} = \lambda \vec{v} \implies \vec{v} = A^{-1}(\lambda \vec{v}) \implies \vec{v} = \lambda (A^{-1} \vec{v}) \implies A^{-1} \vec{v} = \lambda^{-1} \vec{v},$$

so $(\lambda^{-1}, \vec{v})$ is an eigenpair of $A^{-1}$.

2. Matrices $A, B$ are called similar, denotes $A \sim B$ if there exists an invertible matrix $P$ so that $A = PBP^{-1}$. Show the following for any matrices $A, B, C$:

(a) $A \sim A$.

(b) If $A \sim B$, then $B \sim A$.

(c) If $A \sim B$ and $B \sim C$, then $A \sim C$.

Relations $\sim$ with these properties are called equivalence relations, named after everyone’s favorite equivalence relation $=$.
(a) $A = IAI^{-1}$, where $I$ is the identity matrix, so $A \sim A$.

(b) If $A \sim B$, then $A = PBP^{-1}$, so then

$$B = P^{-1}AP = (P^{-1})A(P^{-1})^{-1},$$

which means $B \sim A$.

(c) If $A \sim B$ and $B \sim C$, we have

$$A = PBP^{-1}, \quad B = QCQ^{-1}$$

for some matrices $P, Q$. Note there is no guarantee $P, Q$ are the same, for each similarity relation has its own matrix needed to get from one to the other by $A = PBP^{-1}$. Then,

$$A = PBP^{-1} = PQBQ^{-1}P^{-1} = (PQ)(PQ)^{-1},$$

which means that $A \sim C$.

3. Suppose that $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis of $\mathbb{R}^n$ which are eigenvectors of both $A$ and $B$. That is, say

(i) $(\lambda_1, \vec{v}_1), \ldots, (\lambda_n, \vec{v}_n)$ are eigenpairs of $A$

(ii) $(\rho_1, \vec{v}_1), \ldots, (\rho_n, \vec{v}_n)$ are eigenpairs of $B$.

Show that $AB = BA$.

Method 1: Let $P$ be the matrix with columns $\vec{v}_1, \ldots, \vec{v}_n$ from left to right, $D$ be the diagonal matrix with $\lambda_1, \ldots, \lambda_n$ down the diagonal, and $E$ be the diagonal matrix with $\rho_1, \ldots, \rho_n$ down the diagonal. By diagonalization, we have

$$A = PDP^{-1}, \quad B = PEP^{-1}$$

Thus,

$$AB = (PDP^{-1})(PEP^{-1}) = PDEP^{-1},$$

$$BA = (PEP^{-1})(PDP^{-1}) = PEDP^{-1}.$$  

To multiply two diagonal matrices, we simply multiply entries in corresponding spots along the diagonal, so

$$DE = ED = \begin{bmatrix}
\lambda_1 \rho_1 \\
\lambda_2 \rho_2 \\
\vdots \\
\lambda_n \rho_n
\end{bmatrix}$$

Hence, we conclude

$$AB = PDEP^{-1} = PEDP^{-1} = BA.$$  

Method 2: Consider any $\vec{x} \in \mathbb{R}^n$, which we can write as a linear combination of the basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ for $\mathbb{R}^n$: 

$$\vec{x} = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n.$$  

We have $A\vec{v}_j = \lambda_j \vec{v}_j$ and $B\vec{v}_j = \rho_j \vec{v}_j$ for all $j$. Then,

$$(AB)\vec{x} = A(B(c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n))$$

$$= A(c_1 B\vec{v}_1 + \cdots + c_n B\vec{v}_n)$$

$$= A(c_1 \rho_1 \vec{v}_1 + \cdots + c_n \rho_n \vec{v}_n)$$

$$= c_1 \rho_1 A\vec{v}_1 + \cdots + c_n \rho_n A\vec{v}_n$$

$$= c_1 \rho_1 \lambda_1 \vec{v}_1 + \cdots + c_n \rho_n \lambda_n \vec{v}_n.$$
By the same reasoning,

\[(BA)\vec{x} = c_1\lambda_1\rho_1\vec{v}_1 + \cdots + c_n\lambda_n\rho_n\vec{v}_n.\]

But \(\lambda_i\rho_j = \rho_j\lambda_i\), so we have \(AB\vec{x} = BA\vec{x}\) for all \(\vec{x} \in \mathbb{R}^n\). This means \(AB = BA\).

\[\square\]

4. Let \(B\) be a basis for \(\mathbb{R}^n\), and let \(S, T : \mathbb{R}^n \to \mathbb{R}^n\) be linear maps. Use the Change of Basis Formula to show the following:

(a) \([S \circ T]_B = [S]_B[T]_B\).

(b) If \(T\) is invertible, \([T^{-1}]_B = [T]_B^{-1}\).

(c) Let \(T^k = T \circ T \circ \cdots \circ T\), so \(T\) is done \(k\) times. Then, \([T^k]_B = ([T]_B)^k\) for all positive integers \(k\).

Let \(U\) be the change of basis matrix for basis \(B\). The Change of Basis formula tells us

\([T]_B = U^{-1}[T]U\).

(a) Directly applying the Change of Basis formula,

\([S]_B[T]_B = U^{-1}[S][U]U^{-1}[T]U = U^{-1}[S][T]U = U^{-1}[S \circ T]U = [S \circ T]_B\).

(b) For this part, we will need to be able to take inverse of a product of 3 things. Using the result that \((CD)^{-1} = D^{-1}C^{-1}\), we have

\[(CDE)^{-1} = ((CD)(E))^{-1} = E^{-1}(CD)^{-1} = E^{-1}D^{-1}C^{-1}.\]

Then, applying the Change of Basis formula.

\(([T]_B)^{-1} = (U^{-1}[T]U)^{-1} = U^{-1}[T]^{-1}(U^{-1})^{-1} = U^{-1}[T^{-1}]U = [T^{-1}]_B.\)

(c) We proceed by induction of \(k\). For \(k = 1\), we already know \([T^1]_B = [T]_B = ([T]_B)^1\). Then, assuming that \([T^k]_B = ([T]_B)^k\), we apply 4a) to get

\([T^{k+1}]_B = [T^k \circ T]_B = [T^k]_B[T]_B = ([T]_B)^k[T]_B = ([T]_B)^{k+1}.\)

In other words, if \(A = [T]\), we have

\((U^{-1}AU)^k = U^{-1}A^kU\).

\[\square\]

5. (Extra Credit) Suppose \(T : \mathbb{R}^n \to \mathbb{R}^n\) is diagonalizable, and \(p\) is the characteristic polynomial of \(T\). Show that \(p(T) = O\), the zero matrix.

Note: For example, if the characteristic polynomial of \(T\) is \(\lambda^2 - 6\lambda + 4\), then \(p(T) = T^2 - 6T + 4I\). Also, the above statement is still true when \(T\) is not diagonalizable, but requires more background to prove.

Method 1: Since \(T\) is diagonalizable, there is a basis \(\vec{v}_1, \ldots, \vec{v}_n\) of eigenvectors of \(T\), say with eigenvalues \(\lambda_1, \ldots, \lambda_n\), respectively. This means \(T(\vec{v}_j) = \lambda_j\vec{v}_j\) for all \(j\). This means the characteristic polynomial has \(\lambda_1, \ldots, \lambda_n\) as its \(n\) zeros, which means

\(p(x) = (x - \lambda_1) \cdots (x - \lambda_n).\)
Consider the linear map/matrix
\[ Q = p(T) = (T - \lambda_1 I) \cdots (T - \lambda_n I). \]

First, we show that \( Q \) sends each of \( v_1', \ldots, v_n' \) to \( vec0 \). For \( j = 1, \ldots, n \),
\[
Q(v_j') = (T - \lambda_1 I) \cdots (T - \lambda_j I) \cdots (T - \lambda_n I)v_j' \\
= (T - \lambda_1 I) \cdots (T - \lambda_j I)(T - \lambda_{j+1} I)v_j' \\
= (T - \lambda_1 I) \cdots (T - \lambda_j I)(T(v_j') - \lambda_j v_j') \\
= (T - \lambda_1 I) \cdots (T - \lambda_n I)(\lambda_j v_j' - \lambda_j v_j') \\
= (T - \lambda_1 I) \cdots (T - \lambda_n I)(0) \\
= 0.
\]

Normally, we cannot just rearrange terms in matrix multiplication. However, we can rearrange the different \((T - \lambda I)\) terms above because they multiply like polynomials, and the order of multiplication of polynomials is irrelevant. So, because \((x - 1)(x - 2) = (x - 2)(x - 1)\), \((T - I)(T - 2I) = (T - 2I)(T - I)\).

So, we have shown that \( Q \) sends \( v_1', \ldots, v_n' \) to \( 0 \). Thus, by Proof HW 3, Problem 1, since \( Q \) and \( O \), the zero matrix agree on the basis \( v_1', \ldots, v_n' \), they must agree everywhere, so \( Q = O \). Or, for all \( \vec{x} \in \mathbb{R}^n, \vec{x} = c_1 v_1' + \cdots + c_n v_n' \). Then,
\[
Q(\vec{x}) = Q(c_1 v_1' + \cdots + c_n v_n') = c_1 Q(v_1') + \cdots + c_n Q(v_n') = 0.
\]

**Method 2:** I originally did not realize this method until some students presented it as their solution, so thanks for making me aware of it. Let \( P \) be the matrix with \( v_1', \ldots, v_n' \) as columns, and \( D \) the diagonal matrix of their corresponding eigenvalues \( \lambda_1, \ldots, \lambda_n \). Then,
\[
[T] = PD^{-1}.
\]

The characteristic polynomial of \( T \) has the form
\[
p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,
\]
and satisfies \( p(\lambda_1) = 0, \ldots, p(\lambda_n) = 0 \). Let \( diag(a_1, \ldots, a_n) \) be the matrix with \( a_1, \ldots, a_n \) down the diagonal. Then, we calculate
\[
p([T]) = p(PDP^{-1}) = (PDP^{-1})^n + a_{n-1}(PDP^{-1})^{n-1} + \cdots + a_1(PDP^{-1}) + a_0(I)
\]
Now, by 4c.) \((PDP^{-1})^k = PD^k P^{-1}\). Let \( U = P^{-1} \). Hence,
\[
p([T]) = PD^n P^{-1} + a_{n-1}PD^{n-1}P^{-1} + \cdots + a_1 PDP^{-1} + (Pa_0IP^{-1})
\]
\[
= P(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0I)P^{-1}
\]
\[
= P(diag(\lambda_1^n, \ldots, \lambda_n^n) + a_{n-1} diag(\lambda_1^{n-1}, \ldots, \lambda_n^{n-1}) + \cdots + a_1 diag(\lambda_1, \ldots, \lambda_n) + a_0 diag(1, \ldots, 1)P^{-1}
\]
\[
= P(diag(\lambda_1^n + a_{n-1}\lambda_1^{n-1} + \cdots + a_1 \lambda_1 + a_0, \ldots, \lambda_n^n + a_{n-1}\lambda_n^{n-1} + \cdots + a_1 \lambda_n + a_0))P^{-1}
\]
\[
= P(diag(p(\lambda_1), \ldots, p(\lambda_n)))P^{-1}
\]
\[
= P(diag(0, \ldots, 0))P^{-1}
\]
\[
= POP^{-1}
\]
\[
= O.
\]