Linear Algebra II Lecture 13

Xi Chen

\(^1\)University of Alberta

November 14, 2014
Outline

1. Properties of Eigenvalues and Eigenvectors
2. Diagonalization of Linear Endomorphism
If \( \mathbf{v} \) is an eigenvector of \( T : V \rightarrow V \) corresponding to \( \lambda \), then \( \mathbf{v} \) is an eigenvector of \( T^m \) corresponding to \( \lambda^m \) since

\[
T(\mathbf{v}) = \lambda \mathbf{v} \Rightarrow T^m(\mathbf{v}) = T^{m-1}(T(\mathbf{v})) = T^{m-1}(\lambda \mathbf{v}) = \lambda T^{m-1}(\mathbf{v}) = \lambda T^{m-2}(T(\mathbf{v})) = \ldots = \lambda^m \mathbf{v}.
\]

More generally, if \( f(x) \) is a polynomial in \( x \) and \( \mathbf{v} \) is an eigenvector of \( T : V \rightarrow V \) corresponding to \( \lambda \), then

\[
f(T)(\mathbf{v}) = (a_0 I + a_1 T + \ldots + a_n T^n)(\mathbf{v}) = a_0 \mathbf{v} + a_1 T(\mathbf{v}) + \ldots + a_n T^n(\mathbf{v}) = a_0 \mathbf{v} + a_1 \lambda \mathbf{v} + \ldots + a_n \lambda^n \mathbf{v} = (a_0 + a_1 \lambda + \ldots + a_n \lambda^n)\mathbf{v} = f(\lambda)\mathbf{v}
\]

So \( \mathbf{v} \) is an eigenvector of \( f(T) \) corresponding to \( f(\lambda) \).
Properties of Eigenvalues and Eigenvectors

If \( \mathbf{v} \) is an eigenvector of \( T : V \rightarrow V \) corresponding to \( \lambda \), then \( \mathbf{v} \) is an eigenvector of \( T^m \) corresponding to \( \lambda^m \) since

\[
T(\mathbf{v}) = \lambda \mathbf{v} \implies T^m(\mathbf{v}) = T^{m-1}(T(\mathbf{v})) = T^{m-1}(\lambda \mathbf{v}) = \lambda T^{m-1}(\mathbf{v}) = \lambda T^{m-2}(T(\mathbf{v})) = \ldots = \lambda^m \mathbf{v}.
\]

More generally, if \( f(x) \) is a polynomial in \( x \) and \( \mathbf{v} \) is an eigenvector of \( T : V \rightarrow V \) corresponding to \( \lambda \), then

\[
f(T)(\mathbf{v}) = (a_0 I + a_1 T + \ldots + a_n T^n)(\mathbf{v}) = a_0 \mathbf{v} + a_1 T(\mathbf{v}) + \ldots + a_n T^n(\mathbf{v}) = a_0 \mathbf{v} + a_1 \lambda \mathbf{v} + \ldots + a_n \lambda^n \mathbf{v} = (a_0 + a_1 \lambda + \ldots + a_n \lambda^n) \mathbf{v} = f(\lambda) \mathbf{v}
\]

So \( \mathbf{v} \) is an eigenvector of \( f(T) \) corresponding to \( f(\lambda) \).
Properties of Eigenvalues and Eigenvectors

If \( \mathbf{v} \) is an eigenvector of \( T : V \to V \) corresponding to \( \lambda \), then \( \mathbf{v} \) is an eigenvector of \( T^m \) corresponding to \( \lambda^m \) since

\[
T(\mathbf{v}) = \lambda \mathbf{v} \Rightarrow T^m(\mathbf{v}) = T^{m-1}(T(\mathbf{v})) = T^{m-1}(\lambda \mathbf{v}) = \lambda T^{m-1}(\mathbf{v}) = \lambda T^{m-2}(T(\mathbf{v})) = \ldots = \lambda^m \mathbf{v}.
\]

More generally, if \( f(x) \) is a polynomial in \( x \) and \( \mathbf{v} \) is an eigenvector of \( T : V \to V \) corresponding to \( \lambda \), then

\[
f(T)(\mathbf{v}) = (a_0 I + a_1 T + \ldots + a_n T^n)(\mathbf{v})
= a_0 \mathbf{v} + a_1 T(\mathbf{v}) + \ldots + a_n T^n(\mathbf{v})
= a_0 \mathbf{v} + a_1 \lambda \mathbf{v} + \ldots + a_n \lambda^n \mathbf{v}
= (a_0 + a_1 \lambda + \ldots + a_n \lambda^n) \mathbf{v} = f(\lambda) \mathbf{v}
\]

So \( \mathbf{v} \) is an eigenvector of \( f(T) \) corresponding to \( f(\lambda) \).
Properties of Eigenvalues and Eigenvectors

\[ T(v) = \lambda v \Rightarrow f(T)(v) = f(\lambda)v \]

For example, find all the possible eigenvalues of \( T : V \to V \) satisfying \( T^3 = T \): Let \( v \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then \( v \neq 0, T(v) = \lambda v \) and

\[
(T^3 - T)(v) = (\lambda^3 - \lambda)v \Rightarrow \lambda^3 - \lambda = 0 \Rightarrow \lambda = -1, 0, 1.
\]

For example, find the eigenvalues of \( T : \mathbb{R}[x] \to \mathbb{R}[x] \) given by

\[ T(f(x)) = f(1 - x). \]

Since \( T^2(f(x)) = T(f(1 - x)) = f(x) \), \( T^2 = I \). Let \( v = f(x) \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then

\[
(T^2 - I)(v) = (\lambda^2 - 1)v \Rightarrow \lambda^2 - 1 = 0 \Rightarrow \lambda = -1, 1.
\]

Both 1 and \(-1\) are eigenvalues of \( T \) since \( T(1) = 1 \) and \( T(x - 1/2) = 1/2 - x \).
Properties of Eigenvalues and Eigenvectors

- **Properties of Eigenvalues and Eigenvectors**

\[ T(v) = \lambda v \Rightarrow f(T)(v) = f(\lambda)v \]

For example, find all the possible eigenvalues of \( T : V \rightarrow V \) satisfying \( T^3 = T \): Let \( v \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then \( v \neq 0, \ T(v) = \lambda v \) and

\[ (T^3 - T)(v) = (\lambda^3 - \lambda)v \Rightarrow \lambda^3 - \lambda = 0 \Rightarrow \lambda = -1, 0, 1. \]

For example, find the eigenvalues of \( T : \mathbb{R}[x] \rightarrow \mathbb{R}[x] \) given by

\[ T(f(x)) = f(1 - x). \]

Since \( T^2(f(x)) = T(f(1 - x)) = f(x), \ T^2 = I \). Let \( v = f(x) \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then

\[ (T^2 - I)(v) = (\lambda^2 - 1)v \Rightarrow \lambda^2 - 1 = 0 \Rightarrow \lambda = -1, 1. \]

Both 1 and -1 are eigenvalues of \( T \) since \( T(1) = 1 \) and \( T(x - 1/2) = 1/2 - x \).
Properties of Eigenvalues and Eigenvectors

\[ T(\mathbf{v}) = \lambda \mathbf{v} \Rightarrow f(T)(\mathbf{v}) = f(\lambda)\mathbf{v} \]

For example, find all the possible eigenvalues of \( T : V \rightarrow V \) satisfying \( T^3 = T \): Let \( \mathbf{v} \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then \( \mathbf{v} \neq 0, \ T(\mathbf{v}) = \lambda \mathbf{v} \) and

\[ (T^3 - T)(\mathbf{v}) = (\lambda^3 - \lambda)\mathbf{v} \Rightarrow \lambda^3 - \lambda = 0 \Rightarrow \lambda = -1, 0, 1. \]

For example, find the eigenvalues of \( T : \mathbb{R}[x] \rightarrow \mathbb{R}[x] \) given by

\[ T(f(x)) = f(1 - x). \]

Since \( T^2(f(x)) = T(f(1 - x)) = f(x) \), \( T^2 = I \). Let \( \mathbf{v} = f(x) \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then

\[ (T^2 - I)(\mathbf{v}) = (\lambda^2 - 1)\mathbf{v} \Rightarrow \lambda^2 - 1 = 0 \Rightarrow \lambda = -1, 1. \]

Both 1 and \(-1\) are eigenvalues of \( T \) since \( T(1) = 1 \) and \( T(x - 1/2) = 1/2 - x \).
Properties of Eigenvalues and Eigenvectors

\[ T(v) = \lambda v \Rightarrow f(T)(v) = f(\lambda)v \]

For example, find all the possible eigenvalues of \( T : V \to V \) satisfying \( T^3 = T \): Let \( v \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then \( v \neq 0, \ T(v) = \lambda v \) and

\[ (T^3 - T)(v) = (\lambda^3 - \lambda)v \Rightarrow \lambda^3 - \lambda = 0 \Rightarrow \lambda = -1, 0, 1. \]

For example, find the eigenvalues of \( T : \mathbb{R}[x] \to \mathbb{R}[x] \) given by

\[ T(f(x)) = f(1 - x). \]

Since \( T^2(f(x)) = T(f(1 - x)) = f(x) \), \( T^2 = I \). Let \( v = f(x) \) be an eigenvector of \( T \) corresponding to \( \lambda \). Then

\[ (T^2 - I)(v) = (\lambda^2 - 1)v \Rightarrow \lambda^2 - 1 = 0 \Rightarrow \lambda = -1, 1. \]

Both 1 and \(-1\) are eigenvalues of \( T \) since \( T(1) = 1 \) and \( T(x - 1/2) = 1/2 - x \).
Let $v_1, v_2, \ldots, v_n$ be the eigenvectors of $T : V \rightarrow V$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then $v_1, v_2, \ldots, v_n$ are linearly independent. Otherwise, there are $x_1, x_2, \ldots, x_n \in \mathbb{R}$, not all zero, such that

$$x_1v_1 + x_2v_2 + \ldots + x_nv_n = 0$$

$$\Rightarrow (\lambda_2 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1v_1 + x_2v_2 + \ldots + x_nv_n) = 0$$

$$\Rightarrow (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)\ldots(\lambda_n - \lambda_1)x_1v_1 = 0 \Rightarrow x_1 = 0$$

Similarly, applying $(\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)$ yields

$$(\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1v_1 + x_2v_2 + \ldots + x_nv_n) = 0$$

$$\Rightarrow (\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)\ldots(\lambda_n - \lambda_2)x_2v_2 = 0 \Rightarrow x_2 = 0$$

Applying $\prod_{j \neq i}(\lambda_j I - T)$ yields $x_i = 0$ for $i = 1, 2, \ldots, n$. 
Let \( v_1, v_2, \ldots, v_n \) be the eigenvectors of \( T : V \rightarrow V \) corresponding to distinct eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \). Then \( v_1, v_2, \ldots, v_n \) are linearly independent. Otherwise, there are \( x_1, x_2, \ldots, x_n \in \mathbb{R} \), not all zero, such that

\[
x_1 v_1 + x_2 v_2 + \ldots + x_n v_n = 0
\]

\[
\Rightarrow (\lambda_2 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1 v_1 + x_2 v_2 + \ldots + x_n v_n) = 0
\]

\[
\Rightarrow (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)\ldots(\lambda_n - \lambda_1)x_1 v_1 = 0 \Rightarrow x_1 = 0
\]

Similarly, applying \( (\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T) \) yields

\[
(\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1 v_1 + x_2 v_2 + \ldots + x_n v_n) = 0
\]

\[
\Rightarrow (\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)\ldots(\lambda_n - \lambda_2)x_2 v_2 = 0 \Rightarrow x_2 = 0
\]

Applying \( \prod_{j \neq i} (\lambda_j I - T) \) yields \( x_i = 0 \) for \( i = 1, 2, \ldots, n \).
Properties of Eigenvalues and Eigenvectors

Let \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) be the eigenvectors of \( T : V \rightarrow V \) corresponding to distinct eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \). Then \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) are linearly independent. Otherwise, there are \( x_1, x_2, \ldots, x_n \in \mathbb{R} \), not all zero, such that

\[
\sum_{i=1}^{n} x_i \mathbf{v}_i = 0
\]

\[
\Rightarrow (\lambda_2 I - T)(\lambda_3 I - T)\cdots(\lambda_n I - T)(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n) = 0
\]

\[
\Rightarrow (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)\cdots(\lambda_n - \lambda_1)x_1 \mathbf{v}_1 = 0 \Rightarrow x_1 = 0
\]

Similarly, applying \( (\lambda_1 I - T)(\lambda_3 I - T)\cdots(\lambda_n I - T) \) yields

\[
(\lambda_1 I - T)(\lambda_3 I - T)\cdots(\lambda_n I - T)(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n) = 0
\]

\[
\Rightarrow (\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)\cdots(\lambda_n - \lambda_2)x_2 \mathbf{v}_2 = 0 \Rightarrow x_2 = 0
\]

Applying \( \prod_{j \neq i} (\lambda_j I - T) \) yields \( x_i = 0 \) for \( i = 1, 2, \ldots, n \).
Let \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) be the eigenvectors of \( T : \mathcal{V} \rightarrow \mathcal{V} \) corresponding to distinct eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \). Then \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) are linearly independent. Otherwise, there are \( x_1, x_2, \ldots, x_n \in \mathbb{R} \), not all zero, such that

\[
x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \ldots + x_n \mathbf{v}_n = 0
\]

\[
\Rightarrow (\lambda_2 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \ldots + x_n \mathbf{v}_n) = 0
\]

\[
\Rightarrow (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)\ldots(\lambda_n - \lambda_1)x_1 \mathbf{v}_1 = 0 \Rightarrow x_1 = 0
\]

Similarly, applying \( (\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T) \) yields

\[
(\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \ldots + x_n \mathbf{v}_n) = 0
\]

\[
\Rightarrow (\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)\ldots(\lambda_n - \lambda_2)x_2 \mathbf{v}_2 = 0 \Rightarrow x_2 = 0
\]

Applying \( \prod_{j \neq i}(\lambda_j I - T) \) yields \( x_i = 0 \) for \( i = 1, 2, \ldots, n \).
Properties of Eigenvalues and Eigenvectors

Let \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) be the eigenvectors of \( T : V \to V \) corresponding to distinct eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \). Then \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) are linearly independent. Otherwise, there are \( x_1, x_2, \ldots, x_n \in \mathbb{R} \), not all zero, such that

\[
  x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \ldots + x_n \mathbf{v}_n = 0
\]

\[
  \Rightarrow (\lambda_2 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \ldots + x_n \mathbf{v}_n) = 0
\]

\[
  \Rightarrow (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)\ldots(\lambda_n - \lambda_1)x_1 \mathbf{v}_1 = 0 \Rightarrow x_1 = 0
\]

Similarly, applying \( (\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T) \) yields

\[
  (\lambda_1 I - T)(\lambda_3 I - T)\ldots(\lambda_n I - T)(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \ldots + x_n \mathbf{v}_n) = 0
\]

\[
  \Rightarrow (\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)\ldots(\lambda_n - \lambda_2)x_2 \mathbf{v}_2 = 0 \Rightarrow x_2 = 0
\]

Applying \( \prod_{j \neq i}(\lambda_j I - T) \) yields \( x_i = 0 \) for \( i = 1, 2, \ldots, n \).
Since eigenvectors corresponding to different eigenvalues are linearly independent,

\[ K(\lambda_1 I - T) \cap K(\lambda_2 I - T) = \{0\} \text{ for } \lambda_1 \neq \lambda_2. \]

More generally,

\[ K(\lambda_1 I - T) \cap (K(\lambda_2 I - T) + K(\lambda_3 I - T) + \ldots + K(\lambda_n I - T)) = \{0\} \]

for \( \lambda_1, \lambda_2, \ldots, \lambda_n \) distinct. In other words, if \( B_1 \) is a basis of \( K(\lambda_1 I - T) \), \( B_2 \) is a basis of \( K(\lambda_2 I - T) \), \ldots, and \( B_n \) is a basis of \( K(\lambda_n I - T) \), then \( B_1 \cup B_2 \cup \ldots \cup B_n \) is a linearly independent set (but not necessarily a basis of \( V \)).
Since eigenvectors corresponding to different eigenvalues are linearly independent,

$$K(\lambda_1 I - T) \cap K(\lambda_2 I - T) = \{0\} \text{ for } \lambda_1 \neq \lambda_2.$$ 

More generally,

$$K(\lambda_1 I - T) \cap \left(K(\lambda_2 I - T) + K(\lambda_3 I - T) + \ldots + K(\lambda_n I - T)\right) = \{0\}$$

for $$\lambda_1, \lambda_2, \ldots, \lambda_n$$ distinct. In other words, if $$B_1$$ is a basis of $$K(\lambda_1 I - T)$$, $$B_2$$ is a basis of $$K(\lambda_2 I - T)$$, ..., and $$B_n$$ is a basis of $$K(\lambda_n I - T)$$, then $$B_1 \cup B_2 \cup \ldots \cup B_n$$ is a linearly independent set (but not necessarily a basis of $$V$$).
Since eigenvectors corresponding to different eigenvalues are linearly independent,

\[ K(\lambda_1 I - T) \cap K(\lambda_2 I - T) = \{0\} \text{ for } \lambda_1 \neq \lambda_2. \]

More generally,

\[ K(\lambda_1 I - T) \cap \left( K(\lambda_2 I - T) + K(\lambda_3 I - T) + \ldots + K(\lambda_n I - T) \right) = \{0\} \]

for \( \lambda_1, \lambda_2, \ldots, \lambda_n \) distinct. In other words, if \( B_1 \) is a basis of \( K(\lambda_1 I - T) \), \( B_2 \) is a basis of \( K(\lambda_2 I - T) \), \ldots, and \( B_n \) is a basis of \( K(\lambda_n I - T) \), then \( B_1 \cup B_2 \cup \ldots \cup B_n \) is a linearly independent set (but not necessarily a basis of \( V \)).
(Cayley-Hamilton) Let $V$ be a finite-dimensional vector space and let $T : V \to V$ be a linear transformation with characteristic polynomial $f(x)$. Then

$$f(T) = 0.$$ 

For example, find $T^{2014}$ for $T : \mathbb{R}^2 \to \mathbb{R}^2$ the linear transformation given by $T(x, y) = (x + y, x + y)$. The characteristic polynomial of $T$ is $x^2 - 2x$. Then $T^2 = 2T$ by Cayley-Hamilton. So

$$T^n = T^{n-2}T^2 = T^{n-2}(2T) = 2T^{n-1} = \ldots = 2^{n-1}T$$

and hence

$$T^{2014}(x, y) = 2^{2013}T(x, y) = (2^{2013}(x + y), 2^{2013}(x + y)).$$
(Cayley-Hamilton) Let \( V \) be a finite-dimensional vector space and let \( T : V \to V \) be a linear transformation with characteristic polynomial \( f(x) \). Then
\[
f(T) = 0.
\]

For example, find \( T^{2014} \) for \( T : \mathbb{R}^2 \to \mathbb{R}^2 \) the linear transformation given by \( T(x, y) = (x + y, x + y) \). The characteristic polynomial of \( T \) is \( x^2 - 2x \). Then \( T^2 = 2T \) by Cayley-Hamilton. So
\[
T^n = T^{n-2}T^2 = T^{n-2}(2T) = 2T^{n-1} = \ldots = 2^{n-1}T
\]

and hence
\[
T^{2014}(x, y) = 2^{2013}T(x, y) = (2^{2013}(x + y), 2^{2013}(x + y)).
\]
(Cayley-Hamilton) Let $V$ be a finite-dimensional vector space and let $T : V \rightarrow V$ be a linear transformation with characteristic polynomial $f(x)$. Then

$$f(T) = 0.$$ 

For example, find $T^{2014}$ for $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ the linear transformation given by $T(x, y) = (x + y, x + y)$. The characteristic polynomial of $T$ is $x^2 - 2x$. Then $T^2 = 2T$ by Cayley-Hamilton. So

$$T^n = T^{n-2}T^2 = T^{n-2}(2T) = 2T^{n-1} = \ldots = 2^{n-1}T$$

and hence

$$T^{2014}(x, y) = 2^{2013}T(x, y) = (2^{2013}(x + y), 2^{2013}(x + y)).$$
(Cayley-Hamilton) Let $V$ be a finite-dimensional vector space and let $T : V \to V$ be a linear transformation with characteristic polynomial $f(x)$. Then

$$f(T) = 0.$$ 

For example, find $T^{2014}$ for $T : \mathbb{R}^2 \to \mathbb{R}^2$ the linear transformation given by $T(x, y) = (x + y, x + y)$. The characteristic polynomial of $T$ is $x^2 - 2x$. Then $T^2 = 2T$ by Cayley-Hamilton. So

$$T^n = T^{n-2}T^2 = T^{n-2}(2T) = 2T^{n-1} = ... = 2^{n-1}T$$

and hence

$$T^{2014}(x, y) = 2^{2013}T(x, y) = (2^{2013}(x + y), 2^{2013}(x + y)).$$
(Cayley-Hamilton) Let $V$ be a finite-dimensional vector space and let $T : V \to V$ be a linear transformation with characteristic polynomial $f(x)$. Then

$$f(T) = 0.$$ 

For example, find $T^{2014}$ for $T : \mathbb{R}^2 \to \mathbb{R}^2$ the linear transformation given by $T(x, y) = (x + y, x + y)$. The characteristic polynomial of $T$ is $x^2 - 2x$. Then $T^2 = 2T$ by Cayley-Hamilton. So

$$T^n = T^{n-2} T^2 = T^{n-2} (2T) = 2T^{n-1} = \ldots = 2^{n-1} T$$

and hence

$$T^{2014}(x, y) = 2^{2013} T(x, y) = (2^{2013}(x + y), 2^{2013}(x + y)).$$
Diagonalization of Matrices

A square matrix $A$ is *diagonalizable* if $A$ is similar to a diagonal matrix, i.e., there exists an invertible matrix $P$ such that $PAP^{-1}$ is diagonal.

Diagonalization of Linear Endomorphisms

A linear transformation $T : V \to V$ is *diagonalizable* if there exists a basis $B$ of $V$ such that $[T]_{B \leftarrow B}$ is diagonal. If such $B$ exists, $T$ is *diagonalized* by $B$.

If $[T]_{B \leftarrow B}$ is not diagonal, find another basis $B'$ such that

$$[T]_{B' \leftarrow B'} = P_{B' \leftarrow B} [T]_{B \leftarrow B} P_{B' \leftarrow B}^{-1}$$

is diagonal.

So diagonalizing $T$ is equivalent to diagonalizing $[T]_{B \leftarrow B}$. 
Properties of Eigenvalues and Eigenvectors

Diagonalization of Linear Endomorphism

Diagonalization of Matrices

A square matrix $A$ is *diagonalizable* if $A$ is similar to a diagonal matrix, i.e., there exists an invertible matrix $P$ such that $PAP^{-1}$ is diagonal.

Diagonalization of Linear Endomorphisms

A linear transformation $T : V \rightarrow V$ is *diagonalizable* if there exists a basis $B$ of $V$ such that $[T]_{B \leftarrow B}$ is diagonal. If such $B$ exists, $T$ is diagonalized by $B$.

If $[T]_{B \leftarrow B}$ is not diagonal, find another basis $B'$ such that $[T]_{B' \leftarrow B'} = P_{B' \leftarrow B}[T]_{B \leftarrow B}P_{B' \leftarrow B}^{-1}$ is diagonal.

So diagonalizing $T$ is equivalent to diagonalizing $[T]_{B \leftarrow B}$.
### Diagonalization of Matrices

A square matrix $A$ is *diagonalizable* if $A$ is similar to a diagonal matrix, i.e., there exists an invertible matrix $P$ such that $P A P^{-1}$ is diagonal.

### Diagonalization of Linear Endomorphisms

A linear transformation $T : V \rightarrow V$ is *diagonalizable* if there exists a basis $B$ of $V$ such that $[T]_{B \leftarrow B}$ is diagonal. If such $B$ exists, $T$ is *diagonalized* by $B$.

If $[T]_{B \leftarrow B}$ is not diagonal, find another basis $B'$ such that

$$[T]_{B' \leftarrow B'} = P_{B' \leftarrow B}[T]_{B \leftarrow B}P_{B' \leftarrow B}^{-1}$$

is diagonal.

So diagonalizing $T$ is equivalent to diagonalizing $[T]_{B \leftarrow B}$. 
Non-Diagonalizable Linear Endomorphisms (Matrices)

Not all linear endomorphisms/matrices are diagonalizable!!!

For example, let $T(x, y) = (x + y, y)$ with the corresponding matrix

$$A = [T]_{B \leftarrow B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$ 

If $T$, or equivalently, $A$ is diagonalizable, then

$$PAP^{-1} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

for some $P$. Then $(x - \lambda_1)(x - \lambda_2) = \det(xI - A) = (x - 1)^2$. So $\lambda_1 = \lambda_2 = 1$ and

$$PAP^{-1} = I \Rightarrow A = P^{-1}IP = P^{-1}P = I.$$ 

Contradiction.
Non-Diagonalizable Linear Endomorphisms (Matrices)

Not all linear endomorphisms/matrices are diagonalizable!!!
For example, let \( T(x, y) = (x + y, y) \) with the corresponding matrix

\[
A = [T]_{B \leftrightarrow B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.
\]

If \( T \), or equivalently, \( A \) is diagonalizable, then

\[
PAP^{-1} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}
\]

for some \( P \). Then \( (x - \lambda_1)(x - \lambda_2) = \det(xI - A) = (x - 1)^2 \). So \( \lambda_1 = \lambda_2 = 1 \) and

\[
PAP^{-1} = I \Rightarrow A = P^{-1}IP = P^{-1}P = I.
\]

Contradiction.
Non-Diagonalizable Linear Endomorphisms (Matrices)

Not all linear endomorphisms/matrices are diagonalizable!!! For example, let $T(x, y) = (x + y, y)$ with the corresponding matrix

$$A = [T]_{B^{-}B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$ 

If $T$, or equivalently, $A$ is diagonalizable, then

$$PAP^{-1} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

for some $P$. Then $(x - \lambda_1)(x - \lambda_2) = \det(xI - A) = (x - 1)^2$. So $\lambda_1 = \lambda_2 = 1$ and

$$PAP^{-1} = I \Rightarrow A = P^{-1}IP = P^{-1}P = I.$$ 

Contradiction.
Theorem

Let $V$ be a vector space of dimension $n$. Then a linear transformation $T : V \rightarrow V$ is diagonalizable if and only if $T$ has $n$ linearly independent eigenvectors $v_1, v_2, \ldots, v_n$. In addition, for $B = \{v_1, v_2, \ldots, v_n\}$ if $v_1, v_2, \ldots, v_n$ are linearly independent eigenvectors of $T$ corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_n$. Equivalently, $T$ is diagonalizable if and only if

$$V = K(\lambda_1 I - T) + K(\lambda_2 I - T) + \ldots + K(\lambda_n I - T).$$
In the previous example $T(x, y) = (x + y, y)$ it has one eigenvalue 1 with eigenspace

$$K(I - T) = \{(x, y) : (-y, 0) = 0\} = \{(x, y) : y = 0\} = \text{Span}\{(1, 0)\} \neq \mathbb{R}^2.$$ 

If $T$ has $n$ distinct eigenvalues, then $T$ is diagonalizable. The converse does not hold.

If $T$ is diagonalizable, $f(T)$ is diagonalizable for all polynomials $f(x)$.

If $T$ is diagonalizable and $T$ is bijective, $T^{-1}$ is also diagonalizable.

Every symmetric matrix is diagonalizable. If $[T]_{B \leftarrow B}$ is symmetric, then $T$ is diagonalizable.
In the previous example $T(x, y) = (x + y, y)$ It has one eigenvalue 1 with eigenspace

$$K(I - T) = \{(x, y) : (-y, 0) = 0\} = \{(x, y) : y = 0\} = \text{Span}\{(1, 0)\} \neq \mathbb{R}^2.$$

If $T$ has $n$ distinct eigenvalues, then $T$ is diagonalizable. The converse does not hold.

- If $T$ is diagonalizable, $f(T)$ is diagonalizable for all polynomials $f(x)$.
- If $T$ is diagonalizable and $T$ is bijective, $T^{-1}$ is also diagonalizable.
- Every symmetric matrix is diagonalizable. If $[T]_{B\leftarrow B}$ is symmetric, then $T$ is diagonalizable.
In the previous example $T(x, y) = (x + y, y)$ It has one eigenvalue 1 with eigenspace

$$K(I - T) = \{(x, y) : (-y, 0) = 0\} = \{(x, y) : y = 0\} = \text{Span}\{(1, 0)\} \neq \mathbb{R}^2.$$

If $T$ has $n$ distinct eigenvalues, then $T$ is diagonalizable. The converse does not hold.

- If $T$ is diagonalizable, $f(T)$ is diagonalizable for all polynomials $f(x)$.
- If $T$ is diagonalizable and $T$ is bijective, $T^{-1}$ is also diagonalizable.
- Every symmetric matrix is diagonalizable. If $[T]_{B \leftrightarrow B}$ is symmetric, then $T$ is diagonalizable.
In the previous example $T(x, y) = (x + y, y)$, it has one eigenvalue 1 with eigenspace

$$K(I - T) = \{(x, y) : (-y, 0) = 0\} = \{(x, y) : y = 0\} = \text{Span}\{(1, 0)\} \neq \mathbb{R}^2.$$ 

If $T$ has $n$ distinct eigenvalues, then $T$ is diagonalizable. The converse does not hold.

If $T$ is diagonalizable, $f(T)$ is diagonalizable for all polynomials $f(x)$.

If $T$ is diagonalizable and $T$ is bijective, $T^{-1}$ is also diagonalizable.

Every symmetric matrix is diagonalizable. If $[T]_{B \leftrightarrow B}$ is symmetric, then $T$ is diagonalizable.
In the previous example $T(x, y) = (x + y, y)$ It has one eigenvalue 1 with eigenspace

$$K(I - T) = \{(x, y) : (-y, 0) = 0\} = \{(x, y) : y = 0\}$$
$$= \text{Span}\{(1, 0)\} \neq \mathbb{R}^2.$$

If $T$ has $n$ distinct eigenvalues, then $T$ is diagonalizable. The converse does not hold.

If $T$ is diagonalizable, $f(T)$ is diagonalizable for all polynomials $f(x)$.

If $T$ is diagonalizable and $T$ is bijective, $T^{-1}$ is also diagonalizable.

Every symmetric matrix is diagonalizable. If $[T]_{B\leftarrow B}$ is symmetric, then $T$ is diagonalizable.
In the previous example $T(x, y) = (x + y, y)$ it has one eigenvalue 1 with eigenspace

$$K(I - T) = \{(x, y) : (-y, 0) = 0\} = \{(x, y) : y = 0\} = \text{Span}\{(1, 0)\} \neq \mathbb{R}^2.$$

If $T$ has $n$ distinct eigenvalues, then $T$ is diagonalizable. The converse does not hold.

If $T$ is diagonalizable, $f(T)$ is diagonalizable for all polynomials $f(x)$.

If $T$ is diagonalizable and $T$ is bijective, $T^{-1}$ is also diagonalizable.

Every symmetric matrix is diagonalizable. If $[T]_{B \leftarrow B}$ is symmetric, then $T$ is diagonalizable.
Properties of Eigenvalues and Eigenvectors
Diagonalization of Linear Endomorphism

Upper Triangularization

Theorem

If A is an \( n \times n \) matrix with \( n \) eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \), then A is similar to an upper triangular matrix, i.e., there exists an invertible matrix \( P \) such that

\[
PAP^{-1} = \begin{bmatrix}
\lambda_1 & * & \cdots & * \\
* & \lambda_2 & \cdots & * \\
\vdots & \ddots & \ddots & \vdots \\
* & \cdots & \cdots & \lambda_n
\end{bmatrix}
\]

Proof.

Let \( T : \mathbb{R}^n \rightarrow \mathbb{R}^n \) be the linear transformation given by \( T(v) = Av \). Then \( A = [T]_{B\leftarrow B} \) for the standard basis \( B \).
Theorem

If $A$ is an $n \times n$ matrix with $n$ eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $A$ is similar to an upper triangular matrix, i.e., there exists an invertible matrix $P$ such that

$$PAP^{-1} = \begin{bmatrix}
\lambda_1 & * & \cdots & * \\
& \lambda_2 & \cdots & * \\
& & \ddots & \vdots \\
& & & \lambda_n
\end{bmatrix}$$

Proof.

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the linear transformation given by $T(v) = Av$. Then $A = [T]_{B \leftarrow B}$ for the standard basis $B$. 

Xi Chen

Linear Algebra II Lecture 13
Let $v_1$ be an eigenvector of $A$ corresponding to $\lambda_1$. Then $T(v_1) = \lambda_1 v_1$. We complete $\{v_1\}$ to a basis of $\mathbb{R}^n$:

$$B' = \{v_1, v_2, ..., v_n\}.$$ 

Then

$$[T]_{B' \leftarrow B'} = \begin{bmatrix} [T(v_1)]_{B'} & [T(v_2)]_{B'} & \cdots & [T(v_n)]_{B'} \\ \lambda_1 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \cdots & * \end{bmatrix}$$
Cont.

Let $v_1$ be an eigenvector of $A$ corresponding to $\lambda_1$. Then $T(v_1) = \lambda_1 v_1$. We complete $\{v_1\}$ to a basis of $\mathbb{R}^n$:

$$B' = \{v_1, v_2, ..., v_n\}.$$  

Then

$$[T]_{B' \leftarrow B'} = \begin{bmatrix} [T(v_1)]_{B'} & [T(v_2)]_{B'} & ... & [T(v_n)]_{B'} \\ \lambda_1 & * & * & ... & * \\ 0 & * & * & ... & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & ... & * \end{bmatrix}.$$
Upper Triangularization

Cont.

Then

$$[T]_{B' \leftrightarrow B'} = [T]_{B \leftrightarrow B} P_{B' \leftrightarrow B} P_{B \leftrightarrow B'}$$

$$\Rightarrow PAP^{-1} = \begin{bmatrix} \lambda_1 & * \\ & D \end{bmatrix}$$

By induction on $n$, we can find $Q$ such that $QDQ^{-1}$ is upper triangular. Then

$$\begin{bmatrix} 1 & Q \\ Q & D \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ & D \end{bmatrix} \begin{bmatrix} 1 \\ Q^{-1} \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ & QDQ^{-1} \end{bmatrix}$$

is upper triangular.