Diagonalizability

So far we know that a linear operator T on V or its associated matrix is diagonalizable if and only if there exists an ordered basis $B = \{v_1, \dots, v_n\}$ of eigenvectors of T. However, we don't know yet when a basis of eigenvectors exists. We saw in an example ($A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$) that there are linear operators/matrices which are not diagonalizable.

Theorem: Let T be a linear operator on a vector space V, and let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of T. If $v_1, \dots, v_k$ are eigenvectors of T such that λ_i corresponds to v_i for $1 \leq i \leq k$, then $\{v_1, \dots, v_k\}$ is linearly independent.

Proof: The proof is by induction on k.

For $k = 1$, $v_1 \neq 0$ since it's an eigenvector so $\{v_1\}$ is linearly independent.

Now assume the theorem holds for $k-1$ distinct eigenvalues and let's prove it for k distinct eigenvalues.

Let $\{v_1, \dots v_k\}$ be eigenvectors associated with the distinct eigenvalues $\lambda_1, \dots, \lambda_k$.

Suppose $\{v_1, \dots, v_k\}$ is linearly dependent. Then we have:

$(*)$ $\quad a_1 v_1 + \cdots a_k v_k = 0$; where not all of the $a_i{}'s$ are 0.

Let's apply $T - \lambda_k I$ to both sides of this equations:

$$(T - \lambda_k I)(a_1 v_1 + \cdots a_k v_k) = 0$$

$$a_1 (T - \lambda_k I) v_1 + \cdots a_k (T - \lambda_k I) v_k = 0$$

$$a_1(\lambda_1 - \lambda_k) v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} + a_k(\lambda_k - \lambda_k) = 0.$$

So $$a_1(\lambda_1 - \lambda_k) v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} = 0.$$

But $\{v_1, \ldots, v_{k-1}\}$ is linearly independent so $a_1, \ldots, a_{k-1} = 0$, since $\lambda_i \neq \lambda_k$ if $i \neq k$.

Thus from $(*)$ we get $a_k = 0$.

Hence $\{v_1, \ldots, v_k\}$ is linearly independent.

Corollary: Let T be a linear operator on an n-dimensional vector space V. If T has n distinct eigenvalues then T is diagonalizable.

The converse to the previous theorem is false. That is, if T is diagonalizable it is not true that T must have n distinct eigenvalues. For example, the identity linear operator is diagonal but has only one distinct eigenvalue.

If T is a linear operator on an n-dimensional vector space V, then its characteristic polynomial, $p(\lambda) = \det(A - \lambda I)$, where $[T] = A$, is a polynomial of degree n in λ. In general not every polynomial of degree n can be completely factored into linear factors with real coefficients. For example, $p(\lambda) = \lambda^2 + 1$ can't be factored into linear factors with real coefficients (it can with complex coefficients). If a polynomial of degree n factors completely into linear factors with real coefficients, ie

$$p(\lambda) = c(\lambda - a_1) \cdots (\lambda - a_n)$$

then we say $\boldsymbol{p(\lambda)}$ **splits over $\mathbb{R}$.** Note that the a_i's need not be distinct.

Theorem: The characteristic polynomial of any diagonalizable linear operator splits over $\mathbb{R}$.

Proof: Let T be a diagonalizable linear operator on an n-dimensional vector space V. Then there is an ordered basis for V of eigenvectors $B = \{v_1, \dots, v_n\}$ such that

$$[T]_B = D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

The characteristic polynomial for T is

$$p(\lambda) = \det(D - \lambda I) = det \begin{bmatrix} \lambda_1 - \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n - \lambda \end{bmatrix}$$

$$= (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda).$$

Thus the characteristic polynomial splits over $\mathbb{R}$.

However, the characteristic polynomial of T may split over $\mathbb{R}$ without T being diagonalizable. We saw that with the example $[T] = A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Theorem: Let T be a linear operator on an n-dimensional vector space V. Then T is diagonalizable if and only if:

1. The characteristic polynomial for T splits over $\mathbb{R}$ and
2. For each eigenvalue λ_i of T, the multiplicity of λ_i equals $\dim N(T - \lambda_i I)$.

Ex. Let $A = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}$. Show that A is diagonalizable and find a matrix P such that $D = P^{-1}AP$ is diagonal. Use this diagonal matrix to calculate A^k where k is a positive integer.

We start by finding the characteristic polynomial for A, and solving for its roots.

$$p(\lambda) = \det(A - \lambda I) = det \begin{bmatrix} -\lambda & -2 \\ 1 & 3 - \lambda \end{bmatrix}$$

$$= -\lambda(3 - \lambda) + 2$$

$$= \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1) = 0$$

$$\Longrightarrow \qquad \lambda = 1,2.$$

Since A has two distinct eigenvalues and $\dim(\mathbb{R}^2) = 2$, we know there exist eigenvectors v_1, v_2 for A that form a basis for $\mathbb{R}^2$.

Now let's find the eigenvectors:

For $\lambda = 1$:

$$(A - I)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad -x_1 - 2x_2 = 0 \ \text{ or } \ x_1 = -2x_2.$$

So any vector of the form $< -2\alpha, \alpha >= \alpha < -2,1 >$ is an eigenvector corresponding to $\lambda = 1$. In particular we can use $v_1 =< -2,1 >$.

For $\lambda = 2$:

$$(A - 2I)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -2 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad -2x_1 - 2x_2 = 0 \ \text{ or } \ x_1 = -x_2.$$

So any vector of the form $< -\alpha, \alpha >= \alpha < -1,1 >$ is an eigenvector corresponding to $\lambda = 2$. In particular we can use $v_2 =< -1,1 >$.

The change of basis P will map vectors in terms of the basis $B = \{< -2,1 >, < -1,1 >\}$ into vectors in the standard ordered basis for $\mathbb{R}^2$.

$$P = \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix} \quad \Rightarrow \quad P^{-1} = \frac{1}{\det(P)} \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} = -1 \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$$
$$= \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}.$$

So with respect to the basis B we have:

$$D = P^{-1}AP = \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -2 & -2 \\ 1 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Which is exactly the matrix we would expect given that the eigenvalues of A are 1 and 2.

To calculate A^k for any positive integer k notice the following:

$$D = P^{-1}AP$$
$$PD = AP$$
$$PDP^{-1} = A \quad \Rightarrow \quad (PDP^{-1})^k = A^k.$$

But $\quad (PDP^{-1})^k = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^kP^{-1}.$

So we have: $\quad \boxed{A^k = PD^kP^{-1}.}$

$$A^k = \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^k \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & 2^k \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 2^k & 2^{k+1} \end{bmatrix} = \begin{bmatrix} 2 - 2^k & 2 - 2^{k+1} \\ -1 + 2^k & -1 + 2^{k+1} \end{bmatrix}.$$

Ex. Determine if $T: P_2(\mathbb{R}) \to P_2(\mathbb{R})$ by

$$T\big(p(x)\big) = \big(3p(0) + p'(0)\big) + 3p'(0)x + p''(0)x^2$$

Is diagonalizable.

Let $p(x) = a_0 + a_1 x + a_2 x^2$. Then we have:

$$p(x) = a_0 + a_1 x + a_2 x^2 \quad \Rightarrow \quad p(0) = a_0$$

$$p'(x) = a_1 + 2a_2 x \quad \Rightarrow \quad p'(0) = a_1$$

$$p''(x) = 2a_2 \quad \Rightarrow \quad p''(0) = 2a_2.$$

Thus we can write:

$$T(a_0 + a_1 x + a_2 x^2) = (3a_0 + a_1) + 3a_1 x + 2a_2 x^2.$$

We can find the matrix representation of T with respect to the standard basis $B = \{1, x, x^2\}$ by:

$$T(1) = 3 = 3(1) + 0(x) + 0(x^2) = < 3, 0, 0 >_B$$

$$T(x) = 1 + 3x = 1(1) + 3(x) + 0(x^2) = < 1, 3, 0 >_B$$

$$T(x^2) = 2x^2 = 0(1) + 0(x) + 2(x)^2) = < 0, 0, 2 >_B$$

$$A = [T]_B = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Now let's find the eigenvalues:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 1 & 0 \\ 0 & 3-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (3-\lambda)^2(2-\lambda) = 0.$$

$\lambda = 3$ is a double eigenvalue

$\lambda = 2$ is an eigenvalue.

$\lambda_1 = \lambda_2 = 3$: $\quad A - 3I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ has rank 2.

Notice that $A - 3I: \mathbb{R}^3 \to \mathbb{R}^3$ and $Rank(A - 3I) = 2$. Thus we have:

$$\dim N(A - 3I) + Rank(A - 3I) = \dim(\mathbb{R}^3)$$

$$\dim N(A - 3I) + 2 = 3$$

$\Rightarrow \quad \dim N(A - 3I) = 1 \neq$ multiplicity of λ_1 which is 2.

Thus T is not diagonalizable.

Note: You could also find the eigenspace of $\lambda_1 = \{\alpha < 0,1,0 > | \ \alpha \in \mathbb{R}\}$ which has dimension 1 which is not equal to the multiplicity of λ_1.

Ex. Find the values of k for which $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & k \end{bmatrix}$ is not diagonalizable.

Let's start by finding $p(\lambda) = \det(A - \lambda I)$ and seeing if it splits over $\mathbb{R}$. If it does we can then see if A has any multiple eigenvalues (if it doesn't then A will automatically be diagonalizable). We can then test the multiple eigenvalues to see if $\dim\big(N(A - \lambda I)\big)$ =multiplicity of λ.

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 0 & 0 & k-\lambda \end{vmatrix}.$$ Using the bottom row we get:

$$= (k - \lambda) \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix}$$

$$= (k - \lambda)[(1 - \lambda)^2 - 1] = (k - \lambda)(\lambda)(\lambda - 2).$$

So $p(\lambda)$ splits over $\mathbb{R}$ and the eigenvalues are $\lambda = k, 0,2$.

So the only values of k that will give us any multiple eigenvalues would be $k = 0,2$.

Thus if $k = 0$ in the original matrix then $\lambda = 0$ would be a double eigenvalue.

If $k = 2$ in the original matrix then $\lambda = 2$ would be a double eigenvalue.

Now check to see if the original matrix is diagonalizable for $k = 0$ and/or $k = 2$.

If $k = 0$, then $\lambda = 0$ is a double eigenvalue and

$$A - 0I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow Rank(A) = 1.$$

Thus we have: $\dim\big(N(A-0I)\big) + Rank(A) = \dim(\mathbb{R}^3) = 3$

$$\dim\big(N(A-0I)\big) + \quad 1 \quad = 3$$

$$\Rightarrow \quad \dim\big(N(A-0I)\big) = 2 = \text{multiplicity of } \lambda = 0.$$

Since If $k = 0$, then $\lambda = 0$ is the only multiple eigenvalue, A is diagonalizable.

If $k = 2$, then $\lambda = 2$ is a double eigenvalue and

$$A - 2I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \qquad Rank(A) = 2.$$

Thus we have: $\dim\big(N(A-2I)\big) + Rank(A) = \dim(\mathbb{R}^3) = 3$

$$\dim\big(N(A-0I)\big) + \quad 2 \quad = 3$$

$$\Rightarrow \quad \dim\big(N(A-0I)\big) = 1 \neq \text{multiplicity of } \lambda = 2.$$

Thus if $k = 2$, A is not diagonalizable.