Lecture no 5: EIGENVALUES AND EIGENVECTORS; DIAGONALIZING A MATRIX [1]

1. Determining Eigenvalues and Eigenvectors

Suppose we have a finite dimensional vector space V and a linear operator $F: V \rightarrow V$

If V has dimension n, then F can be expressed as multiplication by an $n \times n$ matrix, but the matrix will depend on the basis for V.

If the scalar field is $K \in \mathbf{C}$, there will be certain "distinguished" vectors, that we call eigenvectors.

If the scalar field is $K \in \mathbf{R}$, this is a possibility, but not a certainty.

We begin our investigation where the scalar field is $\mathbf{R}$ and the vector space is $\mathbf{R}^n$.

Definition. Let V be the vector space $\mathbf{R}^n$ and let A be an $n \times n$ matrix. A nonzero vector $v \in V$ is an eigenvector of A with eigenvalue λ if

$$A v = \lambda v.$$

Note. zero vector 0 cannot be an eigenvector, but 0 can be an eigenvalue. Also, if zero is an eigenvalue for A then A is not a one-to-one function.

The effect of A on an eigenvector depends on the value of λ.

Value of λ	Effect of A on the Eigenvector $\hat{v}$
$0<\lambda<1$	Direction of $\hat{v}$ is unchanged. Length of $\hat{v}$ is decreased.
$\lambda=1$	$\hat{v}$ is unchanged.
$1<\lambda$	Direction of $\hat{v}$ is unchanged. Length of $\hat{v}$ is increased.
$\lambda=0$	$\hat{v}$ becomes the zero vector.
$-1<\lambda<0$	Direction of $\hat{v}$ is reversed. Length of $\hat{v}$ is decreased.
$\lambda=-1$	Direction of $\hat{v}$ is reversed. Length of $\hat{v}$ is decreased.
$\lambda<-1$	Direction of $\hat{v}$ is reversed. Length of $\hat{v}$ is increased.

[1]

7	Л	2	Зв'язок між матрицями лінійного оператора в різних базисах. Власні вектори і власні значення лінійного оператора.
8	ПР	2	Побудова матриці лінійного оператора. Пошук власних векторів і власних значень лінійного оператора.

Our first task is given the matrix A, find the eigenvalues and eigenvectors of A. Consider

$$A\,v = \lambda\,v,\ v \neq 0$$

Then

$$A\,v - \lambda\,v = 0$$

We replace $\lambda\,v$ by $I\lambda\,v$ in the first to get

$$A\,v - I\lambda\,v = 0$$

This is done so that factoring out the v will be legitimate. The expression $(A - \lambda I)$ makes sense, but $(A - \lambda)$ does not.

Now

$$A\,v - I\lambda\,v = (A - \lambda I)v = 0$$

so we have the nonzero vector v in the null space of $(A - \lambda I)$.

This occurs if and only if $(A - \lambda I)$ is not invertible, which occurs if and only if the determinant of $(A - \lambda I)$ is 0.

Thus, we have the following result.

Theorem 1: The eigenvalues of the matrix A are the values of λ for which $\det(A - \lambda I) = 0$.

Example Find the eigenvalues for

$$A = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}.$$

We have

$$A - \lambda I = \begin{pmatrix} 2-\lambda & 1 \\ 0 & -1-\lambda \end{pmatrix}$$

$$det(A - \lambda I) = (2-\lambda)(-1-\lambda) - 0 = \lambda^2 - \lambda - 2.$$

The eigenvalues of this particular matrix are the values of λ for which

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

which are $\lambda = 2$ and $\lambda = -1$.

The expression $\det(A - \lambda I)$, when expanded, is called the *characteristic polynomial* of A, which is denoted $PA(\lambda)$. If A is an $n \times n$ matrix, then the characteristic polynomial of A will have degree n. Expanding $\det(A - \lambda I)$ and finding the values for which $\det(A - \lambda I) = 0$ (i.e., finding the eigenvalues) can be computationally challenging and should be done using a computer in most cases. In the real numbers, the characteristic polynomial may not factor completely, but in the complex numbers, we have the following result.

Theorem 2: In the complex numbers, the characteristic polynomial factors into linear factors and

$$PA(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n)$$

where λ_i are eigenvalues repeated according to their multiplicity.

Example The matrix that rotates a vector through an angle θ in two dimensions is

$$A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

If $\theta = 90^\circ$, then there will not be an eigenvector. (What is the geometric reason there is not?)

To see what happens in this case, note that

$$A(90^\circ) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

so

$$A(90^\circ) - \lambda I = \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix}$$

and

$$det\left[A(90^\circ) - \lambda I\right] = \lambda^2 + 1.$$

Thus, no eigenvalues of $A(90^\circ)$ exist in the real numbers.

This example shows an important difference between the cases when the scalar field is **R** and when the scalar field is **C**. If the scalar field is **C**, the characteristic polynomial will always factor into linear factors.

2. Finding the Eigenvectors after the Eigenvalues Have Been Found

Suppose that we have found the eigenvalues for a matrix A. The next task is to find the eigenvectors associated with each eigenvalue. Suppose that λ_1 is an eigenvalue for A and we want to find the associated eigenvectors. That is, we want to find the nonzero vectors v for which

$$A\,v = \lambda_1\, v \text{ or } (A - \lambda_1 I)v = 0$$

This is exactly the null space of $(A - \lambda_1 I)$ except for the zero vector.

Definition. If A is an $n \times n$ matrix with eigenvalue λ, the eigenspace of λ is the nullspace of $A - \lambda I$. We recall how we found the null space of a matrix B.

Example : Suppose

$$B = \begin{pmatrix} 1 & -1 & 3 \\ -10 & 8 & 8 \\ 14 & -12 & 4 \end{pmatrix}$$

and we want to solve $B\hat{x} = \hat{0}$.

When B is row reduced, the result is

$$\begin{pmatrix} 1 & 0 & -16 \\ 0 & 1 & -19 \\ 0 & 0 & 0 \end{pmatrix}.$$

If we denote the variables as x_1, x_2, x_3, then we have the equations

$$\begin{aligned} x_1 - 16x_3 &= 0, \quad \text{or} \quad x_1 = 16x_3 \\ x_2 - 19x_3 &= 0, \quad \text{or} \quad x_2 = 19x_3 \end{aligned}$$

So x_3 is the free variable and if $x_3 = t$, then $x_1 = 16t$ and $x_2 = 19t$.

Thus, a vector in the null space of B is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 16t \\ 19t \\ t \end{pmatrix} = t\begin{pmatrix} 16 \\ 19 \\ 1 \end{pmatrix}$$

and

$$\left\{ \begin{pmatrix} 16 \\ 19 \\ 1 \end{pmatrix} \right\}$$

is a basis for the null space of B.

We apply this technique to find a basis for each eigenspace. We continue with the example

$$A = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}.$$

To find a basis for the eigenspace for the eigenvalue $\lambda = 2$, we find

$$A - 2I = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -3 \end{pmatrix}.$$

When $A - 2I$ is row reduced, the result is

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

So $x_2 = 0$, and x_1 is the free variable. This gives the eigenvector

$$\begin{pmatrix} x_1 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Putting this in standard form, we have the eigenvectors for $\lambda = 2$ are

$$t\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and a basis for the eigenspace for $\lambda = 2$ is

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$$

If we repeat this procedure for $\lambda = -1$, we get that a basis for the eigenspace for $\lambda = -1$ is

$$\left\{ \begin{pmatrix} 1 \\ -3 \end{pmatrix} \right\}.$$

Theorem 3: A set of eigenvectors, each of which has a different eigenvalue from the others, is a linearly independent set.

Proof We give the proof in the case of three vectors and leave the general case as an exercise. Suppose that v_1, v_2, and v_3 are eigenvectors for A with distinct eigenvalues λ_1, λ_2, λ_3 and suppose that

$$c_1v_1 + c_2v_2 + \ldots + c_3v_3 = 0$$

We will show that each $c_i = 0$

Multiply this equation by $(A - \lambda_3 I)(A - \lambda_2 I)$. Note that

$$(A - \lambda_2 I)\hat{v}_1 = A\hat{v}_1 - \lambda_2\hat{v}_1 = \lambda_1\hat{v}_1 - \lambda_2\hat{v}_1 = (\lambda_1 - \lambda_2)\hat{v}_1$$

so

$$\left(A-\lambda_3 I\right)\left(A-\lambda_2 I\right)\hat{v}_1=\left(A-\lambda_3 I\right)\left[\left(A-\lambda_2 I\right)\hat{v}_1\right]=\left(A-\lambda_3 I\right)\left(\lambda_1-\lambda_2\right)\hat{v}_1$$
$$=\left(\boldsymbol{\lambda}_1-\lambda_2\right)\left(A-\lambda_3 I\right)\hat{v}_1=\left(\lambda_1-\lambda_2\right)\left(\lambda_1-\lambda_3\right)\hat{v}_1.$$

Similarly,

$$\left(A-\lambda_2 I\right)\hat{v}_2=A\hat{v}_2-\lambda_2\hat{v}_2=\lambda_2\hat{v}_2-\lambda_2\hat{v}_2=\hat{0}$$

so

$$\left(A-\lambda_3 I\right)\left(A-\lambda_2 I\right)\hat{v}_2=\left(A-\lambda_3 I\right)\hat{0}=\hat{0}.$$

Finally,

$$\left(A-\lambda_2 I\right)\hat{v}_3=A\hat{v}_3-\lambda_2\hat{v}_3=\lambda_3\hat{v}_3-\lambda_2\hat{v}_3=\left(\lambda_3-\lambda_2\right)\hat{v}_3$$

so

$$\left(A-\lambda_3 I\right)\left(A-\lambda_2 I\right)\hat{v}_3=\left(A-\lambda_3 I\right)\left(\lambda_3-\lambda_2\right)\hat{v}_3=\left(\lambda_3-\lambda_2\right)\left[\left(A-\lambda_3 I\right)\hat{v}_3\right]$$
$$=\left(\boldsymbol{\lambda}_3-\lambda_2\right)\left(\lambda_3-\lambda_3\right)\hat{v}_3=\hat{0}.$$

Thus,

$$\left(A-\lambda_3 I\right)\left(A-\lambda_2 I\right)\left[c_1\hat{v}_1+c_2\hat{v}_2+c_3\hat{v}_3\right]=c_1\left(\lambda_1-\lambda_2\right)\left(\lambda_1-\lambda_3\right)\hat{v}_1+c_2\hat{0}+c_3\hat{0}$$
$$=c_1\left(\lambda_1-\lambda_2\right)\left(\lambda_1-\lambda_3\right)\hat{v}_1=\hat{0}.$$

So,

$$c_1\left(\lambda_1-\lambda_2\right)\left(\lambda_1-\lambda_3\right)\hat{v}_1=\hat{0}$$

and since $(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)\neq 0$ and $\hat{v}_1\neq\hat{0}$, it must be that $c_1=0$.

Similarly, multiplying that equation by $(A-\lambda_3 I)(A-\lambda_1 I)$ yields $c_2 = 0$ and multiplying that equation by $(A-\lambda_2 I)(A-\lambda_1 I)$ yields $c_3 = 0$. ☑

Consequences

(1) If A is a matrix with eigenvector $\hat{v}$ whose eigenvalue is λ, then

$$A^n\hat{v}=\lambda^n\hat{v}.$$

(2) If A is a matrix with eigenvector $\hat{v}$ whose eigenvalue is λ, then for any scalar $\alpha, \alpha\hat{v}$ is also an eigenvector of A with eigenvalue λ.

This is because

$$A\left(\alpha\hat{v}\right)=\alpha\left(A\hat{v}\right)=\alpha\left(\lambda\hat{v}\right)=\lambda\left(\alpha\hat{v}\right).$$

Note. It is important to realize that a single eigenvalue may have more than one linearly independent eigenvector since the null space of $(A - \lambda I)$ may have dimension greater than one.

Example : Consider the two matrices

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

We first find the eigenvalues and eigenvectors for A.

The characteristic polynomial for A is $(2-\lambda)^2$, so $\lambda = 2$ is the only eigenvalue. Now

$$A - 2I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

So there are two free variables, x_1 and x_2, and there are two linearly independent eigenvectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The characteristic polynomial for B is also $(2-\lambda)^2$, so $\lambda = 2$ is again the only eigenvalue. Now

$$B - 2I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

so $x_1 = 0$ and x_2 is free. Thus, there is only one linearly independent eigenvector

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Note. It is important to realize that even if the characteristic polynomials of different matrices are the same, their eigenspaces may be different.

3. Diagonalizing a Matrix

If we have a basis for F^n that consists of eigenvectors of an $n \times n$ matrix A, then the representation of A with respect to that basis is diagonal. The advantages of this from a computational viewpoint are huge.

For example, if

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -.5 \end{pmatrix},$$

then

$$A^n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & (-.5)^n \end{pmatrix}.$$

It is sometimes possible to make sense of a transcendental function of a matrix using Taylor series expansions. We show in the exercises that for certain matrices, expressions such as e^A and $\sin A$ make sense.

The next result is immediate from Theorem 3.

Theorem 4: If the characteristic polynomial of A factors into distinct linear factors, then there is a basis of V that consists of eigenvectors of A.

Definition. The geometric multiplicity of an eigenvalue is the *dimension of the eigenspace* of the eigenvalue.

The algebraic multiplicity of an eigenvalue λ is the exponent of the factor $(x - \lambda)$ in the characteristic polynomial.

A crucial thing to remember in the next discussion is that if A and B are matrices so that $A B$ is defined, then

$$AB = \left[A\hat{b}_1, \ldots, A\hat{b}_k \right],$$

where $A\hat{b}_i$ is the column vector obtained by multiplying the ith column of B on the left by A. We review this in the 2×2 case. If

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

then

$$A\hat{b}_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} \\ a_{21}b_{11} + a_{22}b_{21} \end{pmatrix}$$

$$A\hat{b}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

$$AB=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}.$$

Example : Suppose that A is an $n \times n$ matrix that has n linearly independent eigenvectors $\{v_1, \ldots, v_n\}$ and λ_i is the eigenvalue of v_i. We are not saying that all the λ_i's are distinct. We have

$$A[\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n]=[A\hat{v}_1, A\hat{v}_2, .., A\hat{v}_n]=[\lambda_1\hat{v}_1, \lambda_2\hat{v}_2, \ldots, \lambda_n\hat{v}_n]=[\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n]\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Let P be the matrix whose columns are the eigenvectors $\hat{v}_1, \ldots, \hat{v}_n$ and D the diagonal matrix

$$D=\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Then,

$$AP = PD$$

Note: If Equation above is not clear, consider the 2×2 case where A has eigenvectors

$$\hat{v}_1=\begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix} \quad \hat{v}_2=\begin{pmatrix} v_{12} \\ v_{22} \end{pmatrix}$$

with eigenvalues λ_1 and λ_2, respectively. Then

$$AP=A[\hat{v}_1 \ \hat{v}_2]=[A\hat{v}_1 \ A\hat{v}_2]=[\lambda_1\hat{v}_1 \ \lambda_2\hat{v}_2]=\begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} \\ \lambda_1 v_{21} & \lambda_2 v_{22} \end{pmatrix}$$

and

$$PD=\begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}=\begin{pmatrix} v_{11}\lambda_1+0 & 0+v_{12}\lambda_2 \\ v_{21}\lambda_1+0 & 0+v_{22}\lambda_2 \end{pmatrix}=\begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} \\ \lambda_1 v_{21} & \lambda_2 v_{22} \end{pmatrix}.$$

Typically, we will be given the matrix A and want to find the matrix D. Rearranging the equation gives

$$D = P^{-1}AP$$

We know that P^{-1} exists because $\{v_1, \ldots, v_n\}$ is a basis.

Recapping, we have the following result.

Theorem 5: Let A be a matrix for which there is a basis of eigenvectors $\{v_1, \ldots, v_n\}$. Then

$$P^{-1}AP = D,$$

where P is the matrix whose columns are the eigenvectors of A and D is the diagonal matrix whose diagonal entries are the eigenvalues listed in the same order as the corresponding eigenvectors in P.

Note: Not every square matrix can be diagonalized. Theorem 5 gives the method of accomplishing a diagonalization when it can be done. In most cases, it is not immediately obvious when a matrix can be diagonalized.

There is a special category of matrices that can always be diagonalized and are easily recognized. In $\mathbf{R}^n$, these are the symmetric matrices—matrices for which $A\ T = A$. We will see further that a matrix is symmetric if and only if there is a matrix P for which

$$P = P^T = P^{-1} \qquad \text{and} \qquad P^{-1}AP = \mathrm{D}.$$

Such a matrix P is called an *orthogonal matrix*.

Example : Diagonalize, if possible, the matrix

$$A = \begin{pmatrix} 2 & 0 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

The characteristic polynomial is

$$(2-\lambda)(\lambda-1)(\lambda+1)$$

so there are three eigenvalues: $\lambda = 2, \lambda = 1$, and $\lambda = -1$.

Since the eigenvalues have algebraic dimension one, the matrix can be diagonalized. For

$$\lambda = 2, \text{eigenvector} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}; \quad \lambda = 1, \text{eigenvector} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}; \quad \lambda = -1, \text{eigenvector} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Then, we can take

$$P = \begin{pmatrix} -1 & 0 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

and

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

4. Similar Matrices

If $F : V \to V$ is a linear operator and S is a basis for V, then there is a matrix A for which

$$F(x) = A\,x$$

where the matrix A depends on both F and S. An important distinction between F and A is that F describes how a vector x is changed and is independent of the basis of V, whereas A gives the representation of $F(x)$ in a given coordinate system, and is basis and F dependent.

Definition. If A and B are square matrices, then A is similar to B if there is an invertible matrix P with $B = P^{-1}AP$.

The main result of this section is that two matrices represent the same linear operator with respect to different bases if and only if the matrices are *similar*.

For the purposes of this section, it will be convenient to describe what is meant by an "*equivalence relation*," but that idea will not be used in the sequel. In the present setting, we consider a relation between $n \times n$ matrices. We will say that two such matrices are related if they are *similar*.

A relation is an equivalence relation if three conditions hold. These are
1. A matrix is related to itself. (This is called the reflexive property.)
2. If A is related to B, then B is related to A (symmetric property).
3. If A is related to B and B is related to C, then A is related to C (transitive property).

Theorem 6: Being similar is an equivalence relation on the set of $n \times n$ matrices.

Proof. A is similar to A since $A = I^{-1}AI$.

If A is similar to B then B is similar to A since if $B = P^{-1}AP$, then

$$PBP^{-1} = P\left(P^{-1}AP\right)P^{-1} = \left(PP^{-1}\right)A\left(PP^{-1}\right) = A.$$

Said another way,

$$\left(P^{-1}\right)^{-1}BP^{-1} = A.$$

Also, if A is similar to B and B is similar to C, then A is similar to C because if

$$B = P^{-1}AP \quad \text{and} \quad C = Q^{-1}BQ,$$

then

$$C = Q^{-1}BQ = C = Q^{-1}\left(P^{-1}AP\right)Q = \left(Q^{-1}P^{-1}\right)A(PQ) = (PQ)^{-1}A(PQ).$$

An equivalence relation has the effect of partitioning a set. This means it divides a set into pieces, called equivalence classes, and each element in the set is in exactly one equivalence class. The equivalence classes are determined by the condition that A and B are in the same equivalence class if and only if they are related to each other.

Similar matrices share several characteristics. The next results enumerate some of them.

Theorem 7: Similar matrices have the same characteristic polynomial.

Proof. A is similar to B then there is a matrix B with $B = P^{-1}AP$.

so

$$B - \lambda I = P^{-1}AP - \lambda I = P^{-1}AP - P^{-1}\lambda IP = P^{-1}(A - \lambda I)P$$

and thus

$$det(B - \lambda I) = det\left[P^{-1}(A - \lambda I)P\right] = det\left(P^{-1}\right)\det(A - \lambda I)det(P)$$

$$= \left[det(P)\right]^{-1} det(A - \lambda I)det(P) = det(A - \lambda I).$$

Collorary For similar matrices, the eigenvalues and the algebraic dimensions of their eigenspaces are the same.

While similar matrices have the same characteristic polynomial, two matrices that have the same characteristic polynomial are not necessarily similar.

The next theorem states that two $n \times n$ matrices are similar if and only if they represent the same linear transformation with respect to different bases.

Theorem 8: Let T: : $F^n \rightarrow F^n$ be a linear transformation. Let A be the matrix representation of T with respect to the standard basis. The matrix A is similar to the matrix B if and only if there is a basis of F^n for which B is the representation of T with respect to that basis.

The theorems mentioned earlier say that each linear transformation on a vector space gives rise to an equivalence class of matrices, and that different linear transformations give rise to different equivalence classes of matrices.

We have shown that similar matrices have the same characteristic polynomial, and hence the same eigenvalues and the eigenvalues have the same algebraic multiplicity.

We have now shown that similar matrices have the same geometric multiplicity.

The next example demonstrates that while similar matrices have the same correspondence of eigenvectors and eigenvalues, the representation of the eigenvectors will be different in different bases.

Example : Let $T : \mathbb{R}^3 \to \mathbb{R}^3$. Suppose the representation of T in the standard basis is

$$A = \begin{pmatrix} -2 & 3 & 1 \\ 0 & 1 & 1 \\ -3 & 4 & 1 \end{pmatrix}.$$

The eigenvalues of A are $\lambda = -2, 0,$ and 2. Suppose that $\mathcal{B}$ is the basis

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

so that the change of basis matrix is

$$P_{\mathcal{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Now

$$B = P_{\mathcal{B}}^{-1} A P_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 0 & -3 \\ 0 & 0 & -2 \end{pmatrix}.$$

The eigenvalues of B are also $\lambda = -2, 0,$ and 2.

The eigenvectors of A are

$$\lambda = -2, \quad \hat{v}_{-2} = \begin{pmatrix} 5 \\ -3 \\ 9 \end{pmatrix}; \quad \lambda = 0, \quad \hat{v}_0 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}; \quad \lambda = 2, \quad \hat{v}_2 = \begin{pmatrix} 1 \\ 1 \\ \mathbf{1} \end{pmatrix}.$$

The eigenvectors of B are

$$\lambda=-2,\quad \hat{w}_{-2}=\begin{pmatrix}9\\-12\\8\end{pmatrix}_B;\quad \lambda=0,\quad \hat{w}_0=\begin{pmatrix}-1\\2\\0\end{pmatrix}_B;\quad \lambda=2,\quad \hat{w}_2=\begin{pmatrix}1\\0\\0\end{pmatrix}.$$

Note that the representations of the eigenvectors of B are in the B basis. We show that the eigenvectors are the same; it is just that the representations are different. We check this

$$\hat{w}_{-2}=\begin{pmatrix}9\\-12\\8\end{pmatrix}_B=9\begin{pmatrix}1\\1\\1\end{pmatrix}-12\begin{pmatrix}1\\1\\0\end{pmatrix}+8\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}5\\-3\\9\end{pmatrix}=\hat{v}_{-2}$$

$$\hat{w}_0=\begin{pmatrix}-1\\2\\0\end{pmatrix}_B=-1\begin{pmatrix}1\\1\\1\end{pmatrix}+2\begin{pmatrix}1\\1\\0\end{pmatrix}+0\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\1\\-1\end{pmatrix}=\hat{v}_0$$

$$\hat{\boldsymbol{w}}_2=\begin{pmatrix}1\\0\\0\end{pmatrix}_B=1\begin{pmatrix}1\\1\\1\end{pmatrix}+0\begin{pmatrix}1\\1\\0\end{pmatrix}+0\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\1\\1\end{pmatrix}=\hat{v}_2.$$

If we visualize a vector as an arrow, the eigenvector of a linear operator is a vector that changes only its length (if the eigenvalue is negative, the direction is reversed) when the linear transformation acts on it and the basis in which the eigenvector is represented is immaterial.