5 Eigenvalues and Eigenvectors

5.1 Eigenvectors and Eigenvalues

First note that the title of this chapter and this section are different for some reason. We have studied linear transformations of the form $\bar{x} \mapsto A\bar{x}$ for some matrix A quite extensively. This section will be about those vectors that are, in some sense, unchanged by such a transformation.

Definition An **eigenvector** of an $n \times n$ matrix A is a *nonzero* vector $\bar{x}$ such that $A\bar{x} = \lambda\bar{x}$ for some scalar λ. Similarly, a scalar λ is an **eigenvalue** of A if there is a nontrivial solution $\bar{x}$ of $A\bar{x} = \lambda\bar{x}$, in which case we say $\bar{x}$ is an *eigenvector corresponding to* λ.

In other words, an eigenvector of A is a nonzero vector that gets sent to some multiple of itself under the transformation $\bar{x} \mapsto A\bar{x}$. Checking if some vector is an eigenvector of a matrix is easy; simply multiply the matrix with the vector and check if the result is a multiple of the original vector.

Example Let

$$A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}, \qquad \bar{u} = \begin{bmatrix} -12 \\ 10 \end{bmatrix}, \qquad \bar{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

We determine if $\bar{u}$ and $\bar{v}$ are eigenvectors of A.

$$A\bar{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} -12 \\ 10 \end{bmatrix} = \begin{bmatrix} 48 \\ -40 \end{bmatrix} = -4 \begin{bmatrix} -12 \\ 10 \end{bmatrix},$$

so $\bar{u}$ is an eigenvector of A with corresponding eigenvalue -4. Now we check $\bar{v}$:

$$A\bar{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix},$$

which is *not* a multiple of $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$, so $\bar{v}$ is not an eigenvector of A.

In a spirit similar to that of the preceding example, we can check if some scalar λ is an eigenvalue of A.

Example Here we show that 7 is an eigenvalue of A. First note that

$$A\bar{x} = 7\bar{x} \iff (A - 7I)\bar{x} = \bar{0}.$$

We solve the homogeneous system on the right as usual:

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix},$$

so the general solution has the form

$$x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Every such vector with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$.

In general, λ is an eigenvalue of an $n \times n$ matrix A if and only if

$$(A - \lambda I_n)\bar{x} = \bar{0}$$

has a nontrivial solution. The solution set of this equation is the null space of the matrix $A - \lambda I_n$, which we call the **eigenspace** of A corresponding to λ.

Example Let

$$A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix},$$

which has $\lambda = 2$ as an eigenvalue. We find a basis for the eigenspace corresponding to 2. We reduce the matrix $A - 2I_3$ as follows:

$$\begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -\frac{1}{2} & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

so the general solution to $(A - 2I_3)\bar{x} = \bar{0}$ is

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

A basis for the eigenspace is given by

$$\left\{ \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Notice that the eigenspace is nontrivial if and only if the matrix equation $(A - \lambda I_n)\bar{x} = \bar{0}$ has a free variable (equivalently, there is no free variable if and only if the only solution is the trivial solution $\bar{x} = \bar{0}$). This leads immediately to the following theorem.

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

Proof If A is upper triangular, the matrix $A - \lambda I_n$ looks like

$$\begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} - \lambda & \vdots & & \vdots \\ 0 & 0 & a_{33} - \lambda & & \\ \vdots & \vdots & & \ddots & a_{(n-1)n} \\ 0 & 0 & \ldots & 0 & a_{nn} - \lambda \end{bmatrix}.$$

We see then that there are free variables of $(A - \lambda I_n)\bar{x} = \bar{0}$ if and only if λ is equal to one (or more) of the diagonal entries (there will be no pivot in the corresponding column(s)). The same proof holds if A is lower triangular.

Example The matrix

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

has 3 and 2 as eigenvalues.

It is possible for a matrix to have the scalar 0 as an eigenvalue (even though we do not admit $\bar{0}$ as an eigenvector). If 0 is an eigenvalue, then the corresponding eigenvectors are those $\bar{x}$ such that $A\bar{x} = 0\bar{x} = \bar{0}$. In other words, the eigenspace corresponding to 0 is the solution set of the homogeneous system $A\bar{x} = \bar{0}$.

Theorem If $\bar{v}_1, \ldots, \bar{v}_r$ are eigenvectors that correspond to *distinct* eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\bar{v}_1, \ldots, \bar{v}_r\}$ is linearly independent.

5.2 The Characteristic Equation

We start this section with an example.

Example Let $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$ and suppose we want to find all the eigenvalues of A. First note that since A is not triangular, the trick from the previous section does not apply. However, we can use a determinant to find the eigenvalues. Specifically, we want all λ such that

$$(A - \lambda I_2)\bar{x} = \bar{0}$$

has a nontrivial solution. In order for there to be a nontrivial solution, we need the matrix $A - \lambda I_2$ to have a free variable, which is the same as saying that we need $A - \lambda I_2$ to be non-invertible. In terms of the determinant,

$$\begin{aligned} \det(A - \lambda I_2) &= 0 \\ &\Updownarrow \\ \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} &= 0 \\ &\Updownarrow \\ (2-\lambda)(-6-\lambda) - 9 &= 0 \\ &\Updownarrow \\ \lambda^2 + 4\lambda - 21 &= 0 \\ &\Updownarrow \\ (\lambda - 3)(\lambda + 7) &= 0. \end{aligned}$$

Thus the eigenvalues of A are 3 and -7.

This leads to the following definition.

Definition The scalar equation $\det(A - \lambda I_n) = 0$ is called the **characteristic equation** of A.

Fact A scalar λ is an eigenvalue of the $n \times n$ matrix A if and only if λ satisfies the characteristic equation $\det(A - \lambda I_n) = 0$.

Notice that, when we computed the determinant $\det(A - \lambda I_2)$ in the above example, we ended up with the *polynomial* $\lambda^2 + 4\lambda - 21$. It can be shown that, in general, $\det(A - \lambda I_n)$ is a polynomial in the variable λ, the roots of which are the eigenvalues of A. We call $\det(A - \lambda I_n)$ the **characteristic polynomial** of A.

Example We find the characteristic equation of

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 & 2 \\ 0 & 3 & 1 & 7 & -1 \\ 0 & 0 & 2 & 6 & 5 \\ 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}:$$

$$\begin{aligned} \det(A - \lambda I_5) &= \det \begin{bmatrix} 2-\lambda & 1 & 0 & 0 & 2 \\ 0 & 3-\lambda & 1 & 7 & -1 \\ 0 & 0 & 2-\lambda & 6 & 5 \\ 0 & 0 & 0 & 4-\lambda & 3 \\ 0 & 0 & 0 & 0 & 1-\lambda \end{bmatrix} \\ &= (2-\lambda)(3-\lambda)(2-\lambda)(4-\lambda)(1-\lambda) \\ &= (2-\lambda)^2(3-\lambda)(4-\lambda)(1-\lambda). \end{aligned}$$

Hence the eigenvalues are 1, 2, 3, and 4. We say that 2 is an **eigenvalue of multiplicity 2**, since that is the multiplicity of the factor $(2 - \lambda)$ of the characteristic polynomial.

Example The characteristic equation of a 4×4 matrix is $\lambda^4 - 2\lambda^3 - 8\lambda^2$. We find the eigenvalues and their multiplicities:

$$\begin{aligned} \lambda^4 - 2\lambda^3 - 8\lambda^2 &= \lambda^2(\lambda^2 - 2\lambda - 8) \\ &= \lambda^2(\lambda - 4)(\lambda + 2), \end{aligned}$$

from which we see that 0, 4, and -2 are eigenvalues of A, with 0 having multiplicity 2 and the others having multiplicity 1.

Similarity

Definition Let A and B be $n \times n$ matrices. We say that A **is similar to** B if there is an $n \times n$ invertible matrix P such that

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

or, equivalently, if

$$A = PBP^{-1}.$$

Note It follows easily that if A is similar to B, then B is similar to A.

Theorem If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues with the same multiplicities.

Proof If A is similar to B, then

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

so

$$\begin{aligned} \det(B - \lambda I_n) &= \det(P^{-1}(A - \lambda I_n)P) \\ &= (\det P^{-1})(\det(A - \lambda I_n))(\det P) \\ &= \det(A - \lambda I_n), \end{aligned}$$

which says exactly that A and B have the same characteristic polynomial.

Warnings: (i) Two matrices can have the same eigenvalues but *not* be similar.
(ii) Similarity is *not* the same as row equivalence; row operations usually change the eigenvalues.

Example (to be worked on your own...) Find the characteristic equation and eigenvalues of $A = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

5.3 Diagonalization

Definition An $n \times n$ square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix D, i.e. if $A = PDP^{-1}$ for some diagonal matrix D and invertible matrix P.

Why should such a concept be useful? It allows us to compute powers of A very quickly. Notice that

$$\begin{aligned} A^k &= (PDP^{-1})^k \\ &= \underbrace{(PDP^{-1})(PDP^{-1})\dots(PDP^{-1})}_{k \text{ times}} \\ &= PD^kP^{-1}, \end{aligned}$$

since all the P's in the middle cancel out. Since D is diagonal, it has the form

$$\begin{bmatrix} d_{11} & 0 & 0 & \dots & 0 \\ 0 & d_{22} & 0 & \dots & 0 \\ 0 & 0 & d_{33} & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 0 & d_{nn} \end{bmatrix},$$

and it is easy to see that

$$D^k = \begin{bmatrix} d_{11}^k & 0 & 0 & \dots & 0 \\ 0 & d_{22}^k & 0 & \dots & 0 \\ 0 & 0 & d_{33}^k & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 0 & d_{nn}^k \end{bmatrix}.$$

If k is large, it is much faster to apply such a theorem to compute A^k than it is to multiply A by itself k times. The next natural question to ask is "When is a matrix A diagonalizable?" The following theorem provides an answer.

Theorem An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

Another way to think of this is that A is diagonalizable if and only if there are enough eigenvectors of A to form a basis for $\mathbb{R}^n$. We call such a basis an **eigenvector basis** of $\mathbb{R}^n$.

Now we address the issue of how you actually go about diagonalizing a matrix.

Diagonalizing Matrices To diagonalize an $n \times n$ matrix A, perform the following steps:

1. Find the eigenvalues of A.
If A is triangular, you can simply read the eigenvalues from the diagonal. Otherwise, you should find the zeros of the characteristic polynomial $\det(A - \lambda I_n)$.

2. Find n linearly independent eigenvectors of A.
This is the hard part. For each eigenvalue λ found in step 1, you must find a basis for the corresponding eigenspace, which you do by solving $(A - \lambda I_n)\bar{x} = \bar{0}$ in parametric vector form. Now you hope that you have found n such linearly independent vectors. Otherwise, you stop and conclude that A cannot be diagonalized.

3. Construct P from the vectors found in step 2.
Suppose the vectors found in step 2 are $\{\bar{v}_1, \dots, \bar{v}_n\}$ (at this point the order of these vectors is unimportant). Then you can take P to be

$$P = \begin{bmatrix} \bar{v}_1 & \bar{v}_2 & \dots & \bar{v}_n \end{bmatrix}.$$

4. Construct D from the eigenvectors corresponding to the eigenvalues in part 3.
   Suppose that $\lambda_1, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues corresponding to $\bar{v}_1, \ldots, \bar{v}_n$, respectively (here is where the order matters). Then

$$D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}.$$

Example If possible, diagonalize the matrix

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

1. Find the eigenvalues:

$$\begin{aligned} \det(A - \lambda I_3) &= \det \begin{bmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{bmatrix} \\ &= (1-\lambda)\begin{vmatrix} -5-\lambda & -3 \\ 3 & 1-\lambda \end{vmatrix} - 3\begin{vmatrix} -3 & -3 \\ 3 & 1-\lambda \end{vmatrix} + 3\begin{vmatrix} -3 & -5-\lambda \\ 3 & 3 \end{vmatrix} \\ &= (1-\lambda)[(-5-\lambda)(1-\lambda)+9] - 3[-3+3\lambda+9] + 3[-9+15+3\lambda] \\ &= (1-\lambda)(\lambda^2+4\lambda+4) \\ &= (1-\lambda)(\lambda+2)^2. \end{aligned}$$

   Hence the eigenvalues are 1 and -2.

2. If possible, find three linearly independent eigenvectors. First we solve $(A - I_3)\bar{x} = 0$:

$$\begin{aligned} A - I_3 &= \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

   so

$$\bar{x} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

   It follows that a basis for the eigenspace corresponding to 1 is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

Now we do the same for the equation $(A+2I_3)\bar{x} = \bar{0}$:

$$\begin{aligned} A + 2I_3 &= \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

so

$$\bar{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

A basis for the eigenspace corresponding to -2 is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Now notice that the set we get by combining the bases,

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is itself linearly independent. Hence it is an eigenvector basis of $\mathbb{R}^3$, and we conclude that A is, in fact, diagonalizable.

3. The matrix P is given by

$$P = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

4. The matrix D is given by

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

How can we see if, in fact, the matrices P and D found in the example satisfy the relation $A = PDP^{-1}$? Doing this directly would require we find P^{-1}, which is a pain. The smart way to do this is to note that $A = PDP^{-1}$ if and only if $AP = PD$. Hence you can simply compute AP and PD and check if they are the same. You will find that, in fact, they are.

Theorem An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Notice that this theorem provides a *sufficient*, but not a *necessary*, condition for a matrix to be diagonalizable. For example, the 3×3 matrix A of the preceding example had only 2 distinct eigenvalues, yet was still diagonalizable.

Example We can see easily that the matrix

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

is diagonalizable since it has the three distinct eigenvalues 1, 0, and 3.

If the eigenvalues of a matrix are not distinct, we have the following theorem to tell us if that matrix is diagonalizable.

Theorem Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_p$.

(a) For $1 \leq k \leq p$, the dimension of the eigenspace corresponding to λ_k is less than or equal to the multiplicity of λ_k.

(b) The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n. This occurs if and only if (i) the characteristic polynomial factors *completely* into linear factors, and (ii) the dimension of the eigenspace for each λ_k is equal to the multiplicity of λ_k.

(c) If A is diagonalizable, and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to λ_k, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

Example The matrix

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

has eigenvalues 5 and -3, each having multiplicity 2. It can be shown (check this!) that the bases for the corresponding eigenspaces are

$$\left\{ \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix} \right\}, \qquad \text{and} \qquad \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

By the preceding theorem, the combined set

$$\left\{ \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is linearly independent (and is therefore a basis for $\mathbb{R}^4$).