Math 20
Chapter 5 Eigenvalues and Eigenvectors

1 Eigenvalues and Eigenvectors

1. **Definition:** A scalar λ is called an *eigenvalue* of the $n \times n$ matrix A is there is a nontrivial solution $\mathbf{x}$ of $A\mathbf{x} = \lambda\mathbf{x}$. Such an $\mathbf{x}$ is called an eigenvector *corresponding* to the eigenvalue λ.

2. What does this mean geometrically? Suppose that A is the standard matrix for a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$. Then if $A\mathbf{x} = \lambda\mathbf{x}$, it follows that $T(\mathbf{x}) = \lambda\mathbf{x}$. This means that if $\mathbf{x}$ is an eigenvector of A, then the image of $\mathbf{x}$ under the transformation T is a scalar multiple of $\mathbf{x}$ – and the scalar involved is the corresponding eigenvalue λ. In other words, the image of $\mathbf{x}$ is *parallel* to $\mathbf{x}$.

3. Note that an eigenvector cannot be $\mathbf{0}$, but an eigenvalue can be 0.

4. Suppose that 0 is an eigenvalue of A. What does that say about A? There must be some nontrivial vector $\mathbf{x}$ for which
$$A\mathbf{x} = 0\mathbf{x} = \mathbf{0}$$
which implies that A is not invertible which implies a whole lot of things given our Invertible Matrix Theorem.

5. **Invertible Matrix Theorem Again:** The $n \times n$ matrix A is invertible if and only if 0 is not an eigenvalue of A.

6. **Definition:** The *eigenspace* of the $n \times n$ matrix A corresponding to the eigenvalue λ of A is the set of all eigenvectors of A corresponding to λ.

7. We're not used to analyzing equations like $A\mathbf{x} = \lambda\mathbf{x}$ where the unknown vector $\mathbf{x}$ appears on both sides of the equation. Let's find an equivalent equation in standard form.
$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ A\mathbf{x} - \lambda\mathbf{x} &= \mathbf{0} \\ A\mathbf{x} - \lambda I\mathbf{x} &= \mathbf{0} \\ (A - \lambda I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

8. Thus $\mathbf{x}$ is an eigenvector of A corresponding to the eigenvalue λ if and only if $\mathbf{x}$ and λ satisfy $(A-\lambda I)\mathbf{x} = \mathbf{0}$.

9. It follows that the eigenspace of λ is the null space of the matrix $A - \lambda I$ and hence is a subspace of $\mathbb{R}^n$.

10. Later in Chapter 5, we will find out that it is useful to find a set of linearly independent eigenvectors for a given matrix. The following theorem provides one way of doing so. See page 307 for a proof of this theorem.

11. **Theorem 2:** If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to *distinct* eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

2 Determinants

1. Recall that if λ is an eigenvalue of the $n \times n$ matrix A, then there is a nontrivial solution $\mathbf{x}$ to the equation
$$A\mathbf{x} = \lambda\mathbf{x}$$
or, equivalently, to the equation
$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$
(We call this nontrivial solution $\mathbf{x}$ an eigenvector corresponding to λ.)

2. Note that this second equation has a nontrivial solution if and only if the matrix $A - \lambda I$ is not invertible. Why? If the matrix is not invertible, then it does not have a pivot position in each column (by the Invertible Matrix Theorem) which implies that the homogeneous system has at least one free variable which implies that the homogeneous system has a nontrivial solution. Conversely, if the matrix is invertible, then the only solution is the trivial solution.

3. To find the eigenvalues of A we need a condition on λ that is equivalent to the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$ having a nontrivial solution. This is where determinants come in.

4. We skipped Chapter 3, which is all about determinants, so here's a recap of just what we need to know about them.

5. **Formula:** The determinant of the 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is
$$\det A = ad - bc.$$

6. **Formula:** The determinant of the 3×3 matrix $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ is
$$\begin{aligned} \det A = {} & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}. \end{aligned}$$
See page 191 for a useful way of remembering this formula.

7. **Theorem:** The determinant of an $n \times n$ matrix A is 0 if and only if the matrix A is not invertible.

8. That's useful! We're looking for values of λ for which the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution. This happens if and only if the matrix $A - \lambda I$ is not invertible. This happens if and only if the determinant of $A - \lambda I$ is 0. This leads us to the characteristic equation of A.

3 The Characteristic Equation

1. **Theorem:** A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the *characteristic equation*
$$\det(A - \lambda I) = 0.$$

2. It can be shown that if A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial in the variable λ of degree n. We call this polynomial the *characteristic polynomial* of A.

3. **Example:** Consider the matrix $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$. To find the eigenvalues of A, we must compute $\det(A - \lambda I)$, set this expression equal to 0, and solve for λ. Note that

$$A - \lambda I = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 3-\lambda & 6 & -8 \\ 0 & -\lambda & 6 \\ 0 & 0 & 2-\lambda \end{bmatrix}.$$

Since this is a 3×3 matrix, we can use the formula given above to find its determinant.

$$\begin{aligned} \det(A - \lambda I) &= (3-\lambda)(-\lambda)(2-\lambda) + (6)(6)(0) + (-8)(0)(0) \\ &\quad - (0)(-\lambda)(-8) - (0)(6)(3-\lambda) - (-\lambda)(0)(6) \\ &= -\lambda(3-\lambda)(2-\lambda) \end{aligned}$$

Setting this equal to 0 and solving for λ, we get that $\lambda = 0$, 2, or 3. These are the three eigenvalues of A.

4. Note that A is a triangular matrix. (A triangular matrix has the property that either all of its entries *below* the main diagonal are 0 or all of its entries *above* the main diagonal are 0.) It turned out that the eigenvalues of A were the entries on the main diagonal of A. *This is true for any triangular matrix, but is generally not true for matrices that are not triangular.*

5. **Theorem 1:** The eigenvalues of a triangular matrix are the entries on its main diagonal.

6. In the above example, the characteristic polynomial turned out to be $-\lambda(\lambda-3)(\lambda-2)$. Each of the factors λ, $\lambda - 3$, and $\lambda - 2$ appeared precisely once in this factorization. Suppose the characteristic function had turned out to be $-\lambda(\lambda-3)^2$. In this case, the factor $\lambda - 3$ would appear twice and so we would say that the corresponding eigenvalue, 3, has *multiplicity* 2.

7. **Definition:** In general, the *multiplicity* of an eigenvalue ℓ is the number of times the factor $\lambda - \ell$ appears in the characteristic polynomial.

4 Finding Eigenvectors

1. **Example (Continued):** Let us now find the eigenvectors of the matrix $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$. We have to take each of its three eigenvalues 0, 2, and 3 in turn.

2. To find the eigenvectors corresponding to the eigenvalue 0, we need to solve the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$ where $\lambda = 0$. That is, we need to solve

$$\begin{aligned} (A - \lambda I)\mathbf{x} &= \mathbf{0} \\ (A - 0I)\mathbf{x} &= \mathbf{0} \\ A\mathbf{x} &= \mathbf{0} \\ \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}\mathbf{x} &= \mathbf{0} \end{aligned}$$

Row reducing the augmented matrix, we find that

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}.$$

This tells us that the eigenvectors corresponding to the eigenvalue 0 are precisely the set of scalar multiples of the vector $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$. In other words, the eigenspace corresponding to the eigenvalue 0 is

$$\text{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

3. To find the eigenvectors corresponding to the eigenvalue 2, we need to solve the equation $(A-\lambda I)\mathbf{x} = \mathbf{0}$ where $\lambda = 2$. That is, we need to solve

$$\begin{aligned} (A - \lambda I)\mathbf{x} &= \mathbf{0} \\ (A - 2I)\mathbf{x} &= \mathbf{0} \\ \left(\begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \right) \mathbf{x} &= \mathbf{0} \\ \begin{bmatrix} 1 & 6 & -8 \\ 0 & -2 & 6 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{x} &= \mathbf{0} \end{aligned}$$

Row reducing the augmented matrix, we find that

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -10 \\ 3 \\ 1 \end{bmatrix}.$$

This tells us that the eigenvectors corresponding to the eigenvalue 2 are precisely the set of scalar multiples of the vector $\begin{bmatrix} -10 \\ 3 \\ 1 \end{bmatrix}$. In other words, the eigenspace corresponding to the eigenvalue 2 is

$$\text{Span}\left\{ \begin{bmatrix} -10 \\ 3 \\ 1 \end{bmatrix} \right\}.$$

4. I'll let you find the eigenvectors corresponding to the eigenvalue 3.

5 Similar Matrices

1. **Definition:** The $n \times n$ matrices A and B are said to be *similar* if there is an invertible $n \times n$ matrix P such that $A = PBP^{-1}$.

2. Similar matrices have at least one useful property, as seen in the following theorem. See page 315 for a proof of this theorem.

3. **Theorem 4:** If $n \times n$ matrices are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

4. *Note that if the $n \times n$ matrices A and B are row equivalent, then they are not necessarily similar.* For a simple counterexample, consider the row equivalent matrices $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. If these two matrices were similar, then there would exist an invertible matrix P such that $A = PBP^{-1}$. Since B is the identity matrix, this means that $A = PIP^{-1} = PP^{-1} = I$. Since A is not the identity matrix, we have a contradiction, and so A and B cannot be similar.

5. We can also use Theorem 4 to show that row equivalent matrices are not necessarily similar: Similar matrices have the same eigenvalues but row equivalent matrices often do not have the same eigenvalues. (Imagine scaling a row of a triangular matrix. This would change one of the matrix's diagonal entries which changes its eigenvalues. Thus we would get a row equivalent matrix with different eigenvalues, so the two matrices could not be similar by Theorem 4.)

6 Diagonalization

1. **Definition:** A square matrix A is said to be *diagonalizable* if it is similar to a diagonal matrix. In other words, a diagonal matrix A has the property that there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

2. Why is this useful? Suppose you wanted to find A^3. If A is diagonalizable, then

$$\begin{aligned} A^3 &= (PDP^{-1})^3 = (PDP^{-1})(PDP^{-1})(PDP^{-1}) \\ &= PDP^{-1}PDP^{-1}PDP^{-1} \\ &= PD(PP^{-1})D(PP^{-1})DP^{-1} \\ &= PDDDP^{-1} \\ &= PD^3P^{-1}. \end{aligned}$$

   In general, if $A = PDP^{-1}$, then $A^k = PD^kP^{-1}$.

3. Why is this useful? Because powers of diagonal matrices are relatively easy to compute. For example, if $D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, then

$$D^3 = \begin{bmatrix} 7^3 & 0 & 0 \\ 0 & (-2)^3 & 0 \\ 0 & 0 & 3^3 \end{bmatrix}.$$

   This means that finding A^k involves only two matrix multiplications instead of the k matrix multiplications that would be necessary to multiply A by itself k times.

4. It turns out that an $n \times n$ matrix is diagonalizable if and only it has n linearly independent eigenvectors. That's what the following theorem says. See page 321 for a proof of this theorem.

5. **Theorem 5 (The Diagonalization Theorem):**

   (a) An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

   (b) If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent eigenvectors of A and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are their corresponding eigenvalues, then $A = PDP^{-1}$, where

$$P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$$

   and

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

   (c) If $A = PDP^{-1}$ and D is a diagonal matrix, then the columns of P must be linearly independent eigenvectors of A and the diagonal entries of D must be their corresponding eigenvalues.

6. What can we make of this theorem? If we can find n linearly independent eigenvectors for an $n \times n$ matrix A, then we know the matrix is diagonalizable. Furthermore, we can use those eigenvectors and their corresponding eigenvalues to find the invertible matrix P and diagonal matrix D necessary to show that A is diagonalizable.

7. Theorem 4 told us that similar matrices have the same eigenvalues (with the same multiplicities). So if A is similar to a diagonal matrix D (that is, if A is diagonalizable), then the eigenvalues of D must be the eigenvalues of A. Since D is a diagonal matrix (and hence triangular), the eigenvalues of D must lie on its main diagonal. Since these are the eigenvalues of A as well, the eigenvalues of A must be the entries on the main diagonal of D. This confirms that the choice of D given in the theorem makes sense.

8. See your class notes or Example 3 on page 321 for examples of the Diagonalization Theorem in action.