Lecture 8: Eigenvalues and Eigenvectors

Historic Notes

In the late 1700s Joseph-Louis Lagrange (1736–1813) attempted to prove that the solar system was stable—that is, that the planets would not ever widely deviate from their orbits. Lagrange modeled planetary motion using differential equations. He was assisted in his effort by Pierre-Simon Laplace (1749–1827). Together they reduced the solution of the differential equations to what in actuality was an eigenvalue problem for a matrix of coefficients determined by their knowledge of the planetary orbits. Without having any official notion of matrices, they constructed a quadratic form from the array of coefficients and essentially uncovered the eigenvalues and eigenvectors of the matrix by studying the quadratic form.

The Characteristic Polynomial

Recall that a linear transformation $T : V \to V$ is *diagonalizable* if there is an basis $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for V so that the matrix for T with respect to that basis is diagonal. This means precisely that, for some scalars $\lambda_1, ..., \lambda_n$, we have

$$\begin{aligned} T(\mathbf{v}_1) &= \lambda_1 \mathbf{v}_1 \\ T(\mathbf{v}_2) &= \lambda_2 \mathbf{v}_2 \\ &\vdots \\ T(\mathbf{v}_n) &= \lambda_n \mathbf{v}_n. \end{aligned}$$

An $n \times n$ matrix A is diagonalizable if the associated linear transformation $\mu_A : \mathbb{R}^n \to \mathbb{R}^n$ is diagonalizable. So A is diagonalizable precisely when there is a basis $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for $\mathbb{R}^n$ with the property that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$. We can write these equations in matrix form:

$$A \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

Thus, if we let P be the $n \times n$ matrix whose columns are the vectors $\mathbf{v}_1, ..., \mathbf{v}_n$ and Λ be the $n \times n$ diagonal matrix with diagonal entries $\lambda_1, ..., \lambda_n$, then we have

$$AP = P\Lambda, \quad \text{and so} \quad P^{-1}AP = \Lambda.$$

Definition 1.0.1. *Let $T : V \to V$ be a linear transformation. A* nonzero *vector $\mathbf{v} \in V$ is called an* eigenvector *of T if there is a scalar λ so that $T(\mathbf{v}) = \lambda\mathbf{v}$. The scalar λ is called the associated* eigenvalue *of T.*

This definition leads to a convenient reformulation of diagonalizability:

Proposition 1.0.2. *The linear transformation $T : V \to V$ is diagonalizable if and only if there is a basis for V consisting of eigenvectors of T.*

An important question is how can we find eigenvectors. We most of the time will have a matrix representation for a linear map. Let's say A denote the matrix for T with respect to some basis. We start by observing that the set of eigenvectors with eigenvalue λ, together with the zero vector, forms a subspace.

Lemma 1.0.3. *Let A be an $n \times n$ matrix, and let λ be any scalar. Then*

$$\mathbf{E}(\lambda) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \lambda\mathbf{x}\} = \mathbf{N}(A - \lambda I)$$

is a subspace of $\mathbb{R}^n$. Moreover, $\mathbf{E}(\lambda) \neq \{\mathbf{0}\}$ if and only if λ is an eigenvalue, in which case we call $\mathbf{E}(\lambda)$ the λ-eigenspace of the matrix A.

Proof. The null space $\mathbf{N}(A - \lambda I)$ is a subspace. By definition λ is an eigenvalue if and only if there is a nonzero vector in $\mathbf{E}(\lambda)$ □

The following result is a computational tool for finding eigenvalues.

Proposition 1.0.4. *Let A be an $n \times n$ matrix. Then λ is an eigenvalue of A if and only if* $\det(A - \lambda I) = 0$.

Proof. From above lemma we know that λ is an eigenvalue if and only if the matrix $A - \lambda I$ is singular. The determinant of a singular matrix must be zero. □

Example 1.0.5. *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 3 & 1 \\ -3 & 7 \end{bmatrix}$$

Definition 1.0.6. *Let A be a square matrix. Then $p(t) = p_A(t) = \det(A - tI)$ is called the* characteristic polynomial *of A.*

Lemma 1.0.7. *If A and B are similar matrices, the $p_A(t) = p_B(t)$.*

Proof. Suppose $B = P^{-1}AP$. Then

$$p_B(t) = \det(B - tI) = \det(P^{-1}AP - tI) = \det(P^{-1}(A - tI)P) = \det(A - tI) = p_A(t).$$

□

Diagonalizibility

Theorem 1.0.8. *Let $T : V \to V$ be a linear transformation. Suppose $\mathbf{v}_1, ..., \mathbf{v}_k$ are eigenvectors of T with distinct corresponding eigenvalues $\lambda_1, ..., \lambda_k$. Then $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ is a linearly independent set of vectors.*

Proof. Let m be the largest number between 1 and k so that $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$ is linear independent. We need to show that $m = k$.

By contradiction, suppose $m < k$. Then $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$ is linearly independent and $\{\mathbf{v}_1, ..., \mathbf{v}_m, \mathbf{v}_{m+1}\}$ is linearly dependent. Then we have

$$\mathbf{v}_{m+1} = c_1\mathbf{v}_1 + ... + c_m\mathbf{v}_m$$

for some c_i's. Then using $T\mathbf{v}_i = \lambda_i\mathbf{v}_i$, we can have

$$\mathbf{0} = (T - \lambda_{m+1}I)\mathbf{v}_{m+1} = (T - \lambda_{m+1}I)(c_1\mathbf{v}_1 + ... + c_m\mathbf{v}_m) = c_1(\lambda_1 - \lambda_{m+1})\mathbf{v}_1 + ... + c_m(\lambda_m - \lambda_{m+1})\mathbf{v}_m.$$

Since $\lambda_i - \lambda_{m+1} \neq 0$ for all i, and since $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$ is linearly independent, it has to be that

$$c_1 = c_2 = ... = c_m = 0,$$

contradicting $\mathbf{v}_{m+1} \neq \mathbf{0}$. Thus, $m < k$ is impossible. □

We now arrive at our first result that gives a sufficient condition for a linear transformation to be diagonalizable. Note that we require the eigenvalues to be real numbers, the situation with complex eigenvalues will be discussed later.

Corollary 1.0.9. *Suppose V is an n-dimensional vector space and $T : V \to V$ has n distinct (real) eigenvalues. Then T is diagonalizable.*

Proof. Consider the matrix of T with respect to the basis of eigenvectors. □

The previous condition is "sufficient". There are many diagonalizable matrices with repeated eigenvalues. We discuss two ways in which the hypotheses of Corollary can fail: The characteristic polynomial may have complex roots or it may have repeated roots.

Example 1.0.10. *Consider the matrix*

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

The characteristic polynomial is

$$p(t) = t^2 - (\mathrm{tr} A)t + \det A = t^2 - \sqrt{2}t + 1,$$

whose roots are

$$\lambda = \frac{1+i}{\sqrt{2}} \text{ and } \frac{1-i}{\sqrt{2}}.$$

The geometric meaning of A is to rotate the plane through an angle of $\frac{\pi}{4}$. Thus, it comes as no surprise that A has no real eigenvectors, as there can be no line through the origin that is unchanged after a rotation.

Example 1.0.11. *Consider the matrix*

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

Its characteristic polynomial is $p(t) = t^2 - 4t + 4$, so 2 is a repeated eigenvalue. Now let's find the corresponding eigenvectors:

$$\mathbf{N}(A - 2I) = \mathbf{N}\left(\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}\right) = \mathbf{N}\left(\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\right)$$

is one-dimensional, with basis $\{[1,1]^\top\}$. A cannot be diagonalized since this is the only eigenvector and there cannot be basis of eigenvectors.

An "iff" for diagonalizability*

Definition 1.0.12. *Let λ be an eigenvalue of a linear transformation. The* algebraic multiplicity *of λ is its multiplicity as a root of the characteristic polynomial $p(t)$, i.e., the highest power of $t - \lambda$ dividing $p(t)$. The* geometric multiplicity *of λ is the dimension of the λ-eigenspace $\mathbf{E}(\lambda)$.*

Proposition 1.0.13. *Let λ be an eigenvalue of algebraic multiplicity m and geometric multiplicity d. Then $1 \leq d \leq m$.*

Proof. Suppose λ is an eigenvalue of the linear transformation T. Then $d = \dim \mathbf{E}(\lambda) \geq 1$. Now choose a basis $\mathbf{v}_1, ..., \mathbf{v}_d$ for $\mathbf{E}(\lambda)$ and extend it to a basis $\mathcal{B}$=$\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for V (how?). Then the matrix for T with respect to the basis $\mathcal{B}$ is

$$A = \begin{bmatrix} \lambda I_d & B \\ 0 & C \end{bmatrix}.$$

Then the determinant of $A - tI$ satisfies

$$\det(A - tI) = \det((\lambda - t)I_d) \det(C - tI) = (\lambda - t)^d \det(C - tI).$$

Since the characteristic polynomial does not depend on the choice of basis, and since $(t - \lambda)^m$ is the largest pwoer of $(t - \lambda)$ dividing the characteristic polynomial, it follows that $d \leq m$. $\square$

Theorem 1.0.14. *Let $T : V \to V$ be a linear transformation. Let its distinct eigenvalues be $\lambda_1, ..., \lambda_k$ and assume these are all real numbers. Then T is diagonalizable if and only if the geometric multiplicity, d_i, of each λ_i equals its algebraic multiplicity, m_i.*

Spectral Theorem

We now turn to the study of a large class of diagonalizable matrices, the symmetric matrices. Recall that a square matrix A is symmetric when $A = A^\top$. Start with a general symmetric 2 matrix

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

whose characteristic polynomial is $p(t) = t^2 - (a+c)t + (ac - b^2)$. By the quadratic formula, its eigenvalues are

$$\lambda = \frac{(a+c) + -\sqrt{(a-c)^2 + 4b^2}}{2}.$$

The first thing we notice is that both eigenvalues are real. Moreover, the corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} b \\ \lambda_1 - a \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} \lambda_2 - c \\ b \end{bmatrix}$$

note that

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = b(\lambda_2 - c) + (\lambda_1 - a)b = 0,$$

so the eigenvectors are orthogonal. Since there is an orthogonal basis for $\mathbb{R}^2$ consisting of eigenvectors of A, we have an orthonormal basis consisting of eigenvectors of A. That is, by an appropriate rotation of the usual basis, we obtain a diagonalzing basis for A.

In general, we have the following important result.

Theorem 1.0.15. *Let A be a symmetric $n \times n$ matrix. Then*

1. *The eigenvalues of A are real.*

2. *There is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A. That is, there is an orthogonal matrix Q so that $Q^{-1}AQ = \Lambda$ is diagonal.*

Proof. 1. Let $\lambda = a + bi$ be a eigenvalue of A, and consider the real matrix

$$S = (A - (a+bi)I)(A - (a-bi)I) = (A - aI)^2 + b^2 I.$$

Since

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

it follows that $\det S = 0$. Thus S is singular, and then there is a nonzero vector $\mathbf{x} \in \mathbb{R}^n$ such that $S\mathbf{x} = \mathbf{0}$.

Since $S\mathbf{x} = \mathbf{0}$, then the dot product $S\mathbf{x} \cdot \mathbf{x} = 0$, which gives

$$0 = S\mathbf{x} \cdot \mathbf{x} = \|(A - aI)\mathbf{x}\|^2 + b^2 \|\mathbf{x}\|^2.$$

Then it must be true that $(A - aI)\mathbf{x} = \mathbf{0}$, and $b = 0$. Thus $\lambda = a$ is a real number.

2. Let λ_1 be one of the eigenvalues of A, and choose a unit vector $\mathbf{q}_1$ that is an eigenvector with eigenvalue λ_1. Choose $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ to be any orthonormal basis for $((\mathbf{q}_1))^\perp$. Then the matrix for the linear transformation with respect to the basis $\{\mathbf{q}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ is

$$B = \begin{bmatrix} \lambda_1 & * & \dots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{bmatrix}$$

for some $(n-1) \times (n-1)$ matrix C and some entries $*$. By the change-of-basis formula, we have

$$B = Q^{-1}AQ = Q^\top AQ,$$

since Q is an orthogonal matrix. Therefore,

$$B^\top = (Q^\top AQ)^T = Q^\top A^\top Q = Q^\top AQ = B.$$

Since B is symmetric, we have that the entries $*$ are all 0 and that C is symmetric. This process can be continued and we will end up with a orthonormal basis consisting of eigenvectors of A. ☐

Homework

1. Consider the linear transformation
$$T : \mathcal{M}_{n\times n} \to \mathcal{M}_{n\times n}$$
defined by
$$T(X) = X^\top.$$
Find its eigenvalues and the corresponding eigenspaces. (Hint: Consider the equation $X^\top = \lambda X$)

2. Find the eigenvalues and eigenvectors of the following matrices.
   (a) $\begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}$
   (b) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
   (c) $\begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$

3. Show that if λ is an eigenvalue of the 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and either $b \neq 0$ or $\lambda \neq a$, then $\begin{bmatrix} b \\ \lambda - a \end{bmatrix}$ is a corresponding eigenvector.

4. Suppose A is an $n \times n$ matrix with the property that $A^2 = A$.
   (a) Show that if λ is an eigenvalue of A, then $\lambda = 0$ or $\lambda = 1$.
   (b) Prove that A is diagonalizable.

5. Find orthogonal matrices that diagonalize each of the following symmetric matrices.
   (a) $\begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix}$
   (b) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$

6. Show that if λ is the only eigenvalue of a symmetric matrix A, then $A = \lambda I$.