Lecture 11: Eigenvalues and Eigenvectors

Consider a vector space V and the linear transformation transformation

$$F : V \mapsto V. \tag{1}$$

We say that $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) is an *eigenvalue* of F if there exists a *nonzero* vector $v \in V$ such that

$$F(v) = \lambda v. \tag{2}$$

We call v *eigenvector* of F corresponding to the eigenvalue λ.

Note that, by definition, we are not allowing eigenvectors to be zero, i.e., $v = 0_V$ is not an eigenvector. If we allow $v = 0_V$ to be an eigenvector, then any number λ would be an eigenvalue of F. However, we can have eigenvectors corresponding to zero eigenvalues. In this case the eigenvector belongs to the nullspace of F, since $\lambda = 0 \Rightarrow F(v) = 0_V$.

Next, suppose that V is n-dimensional and let $\mathcal{B}_V = \{u_1, \ldots, u_n\}$ be a basis of V. Denote by $A_{\mathcal{B}_V}$ be the matrix associated with the linear transformation F relative to the basis $\mathcal{B}_V$. We have seen that the coordinates of $F(v)$ relative to the basis $\mathcal{B}_V$ can be expressed as

$$[F(v)]_{\mathcal{B}_V} = A_{\mathcal{B}_V}[v]_{\mathcal{B}_V} \tag{3}$$

where

$$[v]_{\mathcal{B}_V} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{4}$$

are the coordinates of v relative to $\mathcal{B}_V$. We have

$$F(v) = \lambda v \quad \Leftrightarrow \quad A_{\mathcal{B}_V}[v]_{\mathcal{B}_V} = \lambda [v]_{\mathcal{B}_V}. \tag{5}$$

Hence, computing eigenvalues and eigenvectors of matrices is equivalent to compute eigenvalues and eigenvectors of linear transformations between finite-dimensional vector spaces.

Remark: Eigenvalues and eigenvectors can be defined also for linear transformations between infinite-dimensional vector spaces. For example, consider the derivative operator

$$F : C^{(\infty)}(\mathbb{R}) \mapsto C^{(\infty)}(\mathbb{R}),$$
$$f(x) \mapsto \frac{df(x)}{dx}.$$

We have seen that d/dx defines a linear transformation (linear operator) between infinite-dimensional vector spaces. an eigenvector of d/dx corrersponding to an eigenvalue λ has the form $\psi(x) = e^{\lambda x}$. In fact

$$\frac{de^{\lambda x}}{dx} = \lambda e^{\lambda x} \qquad \Rightarrow \qquad \frac{d\psi(x)}{dx} = \lambda \psi(x). \tag{6}$$

Eigenvectors belonging to function spaces are often called *eigenfunctions*.

Eigenvalues of a matrix. Consider a $n \times n$ matrix A with real or complex coefficients. If $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) and $v \in \mathbb{R}^n$ (or $\mathbb{C}^n$) are, respectively, and eigenvalue of A and an eigenvector of A corresponding to λ then

$$Av = \lambda v. \tag{7}$$

Equation (7) is also called *eigenvalue problem* for the matrix A. We have

$$Av = \lambda v \quad \Leftrightarrow \quad (A - \lambda I)v = 0_{\mathbb{R}^n}, \tag{8}$$

Hence, the eigenvector v (which is non-zero by definition) is in the nullspace of the matrix $(A - \lambda I)$. This implies that the matrix $A - \lambda I$ is not injective and therefore not invertible. Equivalently, by using the matrix rank theorem we have that

$$\text{rank}(A - \lambda I) = n - \underbrace{\dim(N(A - \lambda I))}_{\geq 1} < n. \tag{9}$$

This shows that the matrix $(A - \lambda I)$ is not full rank and therefore it is not invertible. A necessary and sufficient condition for $(A - \lambda I)$ to be not invertible is

$$p(\lambda) = \det(A - \lambda I) = 0 \qquad \text{(characterististic equation)}, \tag{10}$$

The polynomial

$$p(\lambda) = \det(A - \lambda I) \tag{11}$$

is known as *characteristic polynomial* associated with the matrix A. The characteristic equation (10) implies that the eigenvalues of a matrix A are roots of the characteristic polynomial $p(\lambda)$.

How many eigenvalues do we have for a given $n \times n$ matrix A? The characteristic polynomial $p(\lambda)$ associated with a $n \times n$ matrix A is a polynomial of degree n with real or complex coefficients (complex coefficients if the matrix A has complex entries). By using the fundamental theorem of algebra (see Lecture 3) we conclude that every $n \times n$ matrix has exactly n complex eigenvalues. Some of such eigenvalues may be repeated, in which case we say that they have “algebraic multiplicity” greater than one. In other words, the multiplicity of an eigenvalue as a root of the characteristic polynomial is called *algebraic multiplicity* the eigenvalue.

If the matrix A is real then the characteristic polynomial $p(\lambda)$ has real coefficients and therefore the roots of $p(\lambda)$ are either real or complex conjugates.

Example 1: Compute the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}. \tag{12}$$

The characteristic polynomial is

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{bmatrix} = -(2 - \lambda)(6 + \lambda) - 9, \tag{13}$$

i.e.,

$$p(\lambda) = \lambda^2 + 4\lambda - 21. \tag{14}$$

The eigenvalues of A are roots of $p(\lambda)$. Setting $p(\lambda) = 0$ yields

$$\lambda_{1,2} = -2 \pm \sqrt{4+21} = -2 \pm 5 \qquad \Rightarrow \qquad \lambda_1 = 3, \quad \lambda_2 = -7. \tag{15}$$

In this case, both eigenvalues have algebraic multiplicity one, i.e., they are simple roots of $p(\lambda)$. The characteristic polynomial can be factored as

$$p(\lambda) = (\lambda - 3)(\lambda + 7), \tag{16}$$

suggesting once again that $\lambda = 3$ and $\lambda = -7$ are simple roots.

Example 2: Compute the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 5 & 1 & -5 \\ 0 & 4 & 3 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{17}$$

In this case we have

$$A - \lambda I = \begin{bmatrix} 2-\lambda & 5 & 1 & -5 \\ 0 & 4-\lambda & 3 & 0 \\ 0 & 0 & 2-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} \tag{18}$$

and

$$p(\lambda) = \det(A - \lambda I) = (2-\lambda)^2(4-\lambda)(1-\lambda). \tag{19}$$

Hence, the matrix A has three eigenvalues:

$$\begin{aligned} \lambda_1 &= 2 \qquad \text{with algebraic multiplicity 2,} \\ \lambda_2 &= 4 \qquad \text{with algebraic multiplicity 1,} \\ \lambda_3 &= 1 \qquad \text{with algebraic multiplicity 1.} \end{aligned}$$

Note that the eigenvalues coincides with the diagonal entries of the matrix A. This is a general fact about upper or or lower triangular matrices, i.e., the eigenvalues of such matrices coincides with the diagonal entries of the matrix. For example, the following matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{20}$$

has two eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$, both with algebraic multiplicity 2.

Example 3: Compute the eigenvalues of the following matrix

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}. \tag{21}$$

The characteristic polynomial is

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 2 \\ -1 & 1-\lambda \end{bmatrix} = -(1-\lambda)^2 + 2, \tag{22}$$

i.e.,

$$p(\lambda) = \lambda^2 - 2\lambda + 3. \tag{23}$$

Hence, the eigenvalues are

$$\lambda_1 = 1 + i\sqrt{2} \qquad \lambda_2 = 1 - i\sqrt{2} \tag{24}$$

Note that λ_1 and λ_2 are complex conjugates eigenvalues. Clearly, for 2×2 matrices with real entries the fundamental theorem of algebra tells us that the eigenvalues are either both real or complex conjugates.

Eigenvectors and eigenspaces. By definition, an eigenvector of a $n \times n$ matrix A is a nonzero vector $v \in \mathbb{R}^n$ such that

$$Av = \lambda v. \tag{25}$$

This means that v is an element of the nullspace of $(A - \lambda I)$ since v is mapped onto the zero of $\mathbb{R}^n$ by $(A - \lambda I)$. We know that such a nullspace is a vector subspace of $\mathbb{R}^n$.

In the context of eigenvalue problems, we call $N(A - \lambda I)$ the *eigenspace* of A corresponding to the eigenvalue λ. The dimension of the eigenspace $N(A - \lambda I)$ is called *geometric multiplicity* of the eigenvalue λ. By definition, an eigenvector cannot be zero and therefore the eigenspace corresponding to each eigenvalue has dimension at least equal to one. The dimension of the eigenspace corresponding to a certain eigenvalue can be computed by using the matrix rank theorem.

Example 4: Compute the eigenspaces of the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} \tag{26}$$

We have seen in a previous example that the eigenvalues of A are $\lambda_1 = 3$ and $\lambda_2 = -7$. Let us compute the eigenspace corresponding to λ_1. To this end, we first compute the dimension of such eigenspace by using the matrix rank theorem

$$\dim(N(A - \lambda_1 I)) = 2 - \text{rank}(A - \lambda_1 I) = 2 - \text{rank}\left(\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}\right) = 2 - 1 = 1 \tag{27}$$

Hence, the eigenspace corresponding to λ_1 has dimension one. Any vector of such an eigenspace is an eigenvector of A corresponding to λ_1. To compute a basis for the eigenspace $N(A - \lambda_1 I)$ consider

$$(A - \lambda_1 I)v = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad -v_1 + 3v_2 = 0 \tag{28}$$

Hence,

$$v = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{29}$$

is a basis for $N(A - \lambda_1 I)$, and an eigenvector of A corresponding to λ_1. All eigenvectors of A corresponding to λ_1 are in the form

$$c \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{with } c \neq 0. \tag{30}$$

Similarly, the eigenspace corresponding to λ_2 has dimension 1 and can be determined by solving the linear system

$$(A - \lambda_2 I)v = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad 3v_1 + v_2 = 0. \tag{31}$$

Hence,

$$v = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \tag{32}$$

is a basis for $N(A - \lambda_2 I)$ and an eigenvector of A corresponding to λ_2. In summary, λ_1 and λ_2 are eigenvalues with algebraic multiplicity one and geometric multiplicity one. Geometric multiplicity one means that the eigenspaces $N(A-\lambda_1 I)$ and $N(A-\lambda_2 I)$ are both one-dimensional. A basis for $N(A - \lambda_1 I)$ and $N(A - \lambda_2 I)$ is given by (29) and (32), respectively.

The following theorem establishes a relationship between the algebraic multiplicity and the geometric multiplicity of an eigenvalue λ.

Theorem 1. Let λ be an eigenvalue of a $n \times n$ matrix A. Denote by s the algebraic multiplicity of λ. Then

$$\dim(N(A - \lambda I)) \leq s. \tag{33}$$

In other words the geometric multiplicity of λ (i.e., the dimension of the associated eigenspace) is always smaller or equal than the algebraic multiplicity).

Of course, if λ is a simple eigenvalue ($s = 1$) then $\dim(N(A - \lambda I)) = 1$, i.e., the eigenspace corresponding to simple eigenvalues is always one-dimensional. If λ has algebraic multiplicity 2, i.e., it is a repeated eigenvalue, then it is possible to have geometric multiplicity equal to one or equal to two. In the latter case the eigenspace is two-dimensional and any vector in such eigenspace (including linear combinations of multiple eigenvectors) is an eigenvector. Let us provide a simple example of a 2×2 matrix with one eigenvalue of algebraic multiplicity two and geometric multiplicity one

Example 5: Consider the following matrix

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}. \tag{34}$$

We know that $\lambda = 2$ is the only eigenvalue and it has algebraic multiplicity two. In fact, the characteristic polynomial is $p(\lambda) = (2 - \lambda)^2$. The geometric multiplicity of $\lambda = 2$ can be calculated by using the matrix rank theorem

$$\dim(N(A - \lambda I)) = 2 - \text{rank}(A - \lambda I) = 2 - \underbrace{\text{rank}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right)}_{=1} = 2 - 1 = 1. \tag{35}$$

Hence, the eigenspace associated with $\lambda = 2$ is one-dimensional. A basis for such an eigenspace is obtained as follows:

$$(A - \lambda I)v = 0_{\mathbb{R}^2} \quad \Leftrightarrow \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad v_2 = 0. \tag{36}$$

We can choose as basis

$$v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{37}$$

Example 6: Compute the eigenvalues and the eigenvectors of the following matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 2 \end{bmatrix}. \tag{38}$$

This is an upper triangular matrix and therefore the eigenvalues coincide with the diagonal entries. Hence we have $\lambda_1 = 2$ with algebraic multiplicity two and $\lambda_2 = 1$ with algebraic multiplicity one.

$$A - \lambda_1 I = \begin{bmatrix} 0 & 1 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 0 \end{bmatrix} \qquad \Leftrightarrow \qquad \dim(N(A - \lambda_1 I)) = 3 - \underbrace{\text{rank}(A - \lambda_1 I)}_{=2} = 1 \tag{39}$$

$$A - \lambda_2 I = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 1 \end{bmatrix} \qquad \Leftrightarrow \qquad \dim(N(A - \lambda_2 I)) = 3 - \underbrace{\text{rank}(A - \lambda_2 I)}_{=2} = 1 \tag{40}$$

Therefore, the dimension of the eigenspaces associated with λ_1 and λ_2 is one. Let us find a basis for such eigenspaces.

$$(A - \lambda_1 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 \text{ arbitrary} \\ v_2 + 3v_3 = 0 \\ -v_2 + 5v_3 = 0 \end{cases} \tag{41}$$

Hence, an eigenvector that spans $N(A - \lambda_1 I)$ is

$$v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{42}$$

Similarly,

$$(A - \lambda_2 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 + v_2 + 3v_3 = 0 \\ v_3 = 0 \end{cases} \tag{43}$$

Hence, an eigenvector that spans $N(A - \lambda_2 I)$ is

$$v = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}. \tag{44}$$

Theorem 2. Eigenvectors corresponding to different eigenvalues are linearly independent.

Proof. Let v_1 and v_2 be two eigenvectors of a matrix $A \in M_{n\times n}$ corresponding to two distinct eigenvalues λ_1 and λ_2. We want to show that

$$x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^n} \quad \Rightarrow \quad x_1 = x_2 = 0. \tag{45}$$

To this end we first multiply the equation above by λ_2 to obtain

$$x_1 \lambda_2 v_1 + x_2 \lambda_2 v_2 = 0_{\mathbb{R}^n} \tag{46}$$

Then we apply the matrix A to $x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^n}$ to obtain

$$x_1 A v_1 + x_2 A v_2 = x_1 \lambda_1 v_1 + x_2 \lambda_2 v_2 = 0_{\mathbb{R}^n} \tag{47}$$

Subtracting equation (46) from equation (47) yields

$$x_1 \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} \underbrace{v_1}_{\neq 0_{\mathbb{R}^n}} = 0_{\mathbb{R}^n} \qquad \Rightarrow \qquad x_1 = 0. \tag{48}$$

Substituting $x_1 = 0$ into $x_1 v_1 + x_2 v_2 = 0_{\mathbb{R}^n}$ yields $x_2 = 0$. Hence v_1 and v_2 are linearly independent.

□

Similarity transformations. Let $A, B \in M_{n\times n}$. We say that A is *similar* to B is there exists an invertible matrix $P \in M_{n\times n}$ such that

$$AP = PB \qquad \Leftrightarrow \qquad A = PBP^{-1} \tag{49}$$

The transformation $B \to PBP^{-1}$ is called *similarity transformation*. An example of similarity transformation is the change of basis transformation.

Theorem 3. Similar matrices have the same eigenvalues.

Proof. Let $A, B \in M_{n\times n}$ be two similar matrices, i.e., $P \in M_{n\times n}$ such that

$$A = PBP^{-1}. \tag{50}$$

Then

$$\det(A - \lambda I) = \det(PBP^{-1} - \lambda PP^{-1}) = \det(P)\det(B - \lambda I)\det(P^{-1}) = \det(B - \lambda I) \tag{51}$$

□

This theorem implies that the eigenvalues of a linear transformation $F : V \mapsto V$ ($\dim(V) = n$) do not depend on the basis we choose to represent F in V. In fact the matrices associated to F relative to different bases of V are related by a similarity transformation.

Diagonalization. Consider a $n \times n$ matrix A. We have seen in Theorem 2 that eigenvectors corresponding to different eigenvalues are linearly independent. Hence, if the algebraic multiplicity of each eigenvalue is equal to the geometric multiplicity then it is possible to construct a basis for $\mathbb{R}^n$ made of eigenvectors of A. Let us organize such n eigenvectors as columns of a matrix P

$$P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}. \tag{52}$$

Clearly,

$$AP = \begin{bmatrix} Av_1 & \cdots & Av_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \underbrace{\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}}_{\Lambda} = P\Lambda, \tag{53}$$

where Λ is a diagonal matrix having the eigenvalues of A (counted with their multiplicity) along the diagonal. Equation (53) shows that if A has n linearly independent eigenvectors then A is similar to a diagonal matrix[1] Λ. The similarity transformation is defined by the matrix P in (52), i.e., the matrix that has the eigenvectors of A as columns.

A corollary of this statement is that matrices with simple eigenvalues are always diagonalizable, since they have n linearly independent eigenvectors. The following theorem summarizes what we just said.

Theorem 4. Let A be a $n \times n$ matrix with eigenvalues $\{\lambda_1, \ldots, \lambda_p\}$ with algebraic multiplicities $\{s_1, \ldots, s_p\}$, respectively. Then A is diagonalizable if and only if

$$\dim(N(A - \lambda_i I)) = s_i \quad \text{for all} \quad i = 1, \ldots, p. \tag{54}$$

Example 7: The matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} \tag{55}$$

is diagonalizable. In fact we have seen that the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -7$ (simple eigenvalues). This implies that the dimension of the associated eigenspace is one for both eigenvalues. The eigenvectors of A are

$$v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad \text{and} \qquad v_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}. \tag{56}$$

Define

$$P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}. \tag{57}$$

It is straightforward to verify that

$$P^{-1} = \frac{1}{10} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix} \tag{58}$$

and

$$A = P\Lambda P^{-1} \qquad \text{or} \qquad \Lambda = P^{-1}AP. \tag{59}$$

[1] In general, we say that a matrix A is *diagonalizable* if there exists an invertible matrix P such that A is similar to a diagonal matrix.

Example 8: The matrix

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \tag{60}$$

is not diagonalizable. In fact the algebraic multiplicity of the eigenvalue $\lambda = 2$ is two, while its geometric multiplicity is one. It is possible to show that there exists a basis made of "generalized eigenvectors" that makes A similar to a matrix J called *Jordan form* of A. In this particular example, the Jordan form of A coincides with A, i.e., A is already in a Jordan form (see the Remark at page 10).

Example 9: Verify that the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \tag{61}$$

is diagonalizable. The matrix is lower-triangular with eigenvalues $\lambda_1 = 1$ (algebraic multiplicity two) and $\lambda_2 = 2$ (algebraic multiplicity one). To verify that A is diagonalizable we just need to check that the geometric multiplicity of $\lambda_1 = 1$ is equal to two. To this end, we use the matrix rank theorem:

$$\dim(N(A - \lambda_1 I)) = 3 - \text{rank}(A - \lambda_1 I) = 3 - \text{rank}\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}\right) = 3 - 1 = 2 \tag{62}$$

This shows that the dimension of the nullspace of $N(A - \lambda_1 I)$, i.e., the dimension of the eigenspace associated with $\lambda_1 = 1$ is two. Let us compute a basis for such an eigenspace. To this end,

$$(A - \lambda_1 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 \text{ arbitrary} \\ v_2 \text{ arbitrary} \\ v_3 = -v_2 \end{cases} \tag{63}$$

Hence, a basis for the eigenspace corresponding to λ_1 is

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}. \tag{64}$$

On the other hand, the eigenspace $N(A - \lambda_2 I)$ is spanned by a vector that can be computed as

$$(A - \lambda_2 I)v = 0_{\mathbb{R}^3} \quad \Rightarrow \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} v_1 = 0 \\ v_2 = 0 \\ v_3 \text{ arbitrary} \end{cases} \tag{65}$$

Therefore a matrix P that diagonalizes A is

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}. \tag{66}$$

Indeed, it can be verified by a direct calculation that

$$\underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{bmatrix}}_{\Lambda} = \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 1\end{bmatrix}}_{P^{-1}} \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 2\end{bmatrix}}_{A} \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 1\end{bmatrix}}_{P}. \tag{67}$$

Remark: It can be shown that the set of eigenvectors of any $n \times n$ matrix A can be complemented to a basis of $\mathbb{R}^n$ (or $\mathbb{C}^n$) by adding a certain number of *generalized eigenvectors*, as many as $s_i - \dim(N(A - \lambda_i I))$ in case the eigenspace $N(A - \lambda_i I)$ has dimension smaller than the algebraic multiplicity of λ_i. For instance, consider the matrix

$$A = \begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix} \tag{68}$$

We know that the eigenspace corresponding to $\lambda = 2$ is one-dimensional with basis

$$v = \begin{bmatrix}1\\ 0\end{bmatrix}. \tag{69}$$

To complement v to a basis of $\mathbb{R}^2$ we can construct another vector w as follows

$$(A - \lambda I)w = v. \tag{70}$$

Clearly, w is in the nullspace of the matrix $(A - \lambda I)^2$. It can be shown that w and v are linearly independent. We obtain.

$$(A - \lambda I)w = v \quad \Leftrightarrow \quad \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}w_1\\ w_2\end{bmatrix} = \begin{bmatrix}1\\ 0\end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} w_1 \text{ arbitrary} \\ w_2 = 1 \end{cases}. \tag{71}$$

At this point we can define

$$P = \begin{bmatrix}v & w\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \qquad \text{(matrix of generalized eigenvectors)}, \tag{72}$$

and apply A to P to obtain

$$AP = \begin{bmatrix}Av & Aw\end{bmatrix} = \begin{bmatrix}v & w\end{bmatrix} \underbrace{\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}}_{J} = PJ \tag{73}$$

Hence, A is similar to a matrix J in a particular form (not diagonal but almost diagonal), known as *Jordan canonical form.* In this particular example, A is already in a Jordan form so the similarity transformation defined by P turns out to be the identity transformation.

We conclude this section with an important theorem characterizing the *spectral properties* (i.e., eigenvalues and eigenvectors) of real symmetric matrices.

Theorem 5 (Spectral theorem for symmetric matrices)**.** If $A \in M_{n\times n}(\mathbb{R})$ is symmetric (i.e., $A = A^T$) then all eigenvalues are real and there exists an orthonormal basis of $\mathbb{R}^n$ made of eigenvectors of A.

Proof. To prove that the eigenvalues are real let us consider the following scalar product in $\mathbb{C}^n$:

$$\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i^* \tag{74}$$

Suppose that u is an eigenvector of A. Then

$$\langle Au, u\rangle = \langle \lambda u, u\rangle = \lambda \langle u, u\rangle \tag{75}$$

On the other hand,

$$\langle u, Au\rangle = \langle u, \lambda u\rangle = \lambda^* \langle u, u\rangle . \tag{76}$$

The matrix A is symmetric. This implies that

$$\langle Au, u\rangle = \langle u, Au\rangle \qquad \Leftrightarrow \qquad \lambda^* = \lambda, \tag{77}$$

i.e., λ is real.

Let us now prove that eigenvectors of A corresponding to different eigenvalues are necessarily orthogonal. To this end, suppose that u_1 and u_2 are eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2. Then

$$\lambda_1 \langle u_1, u_2\rangle = \langle Au_1, u_2\rangle = \langle u_1, Au_2\rangle = \lambda_2 \langle u_1, u_2\rangle . \tag{78}$$

Since $\lambda_1 \neq \lambda_2$ we have that the previous equality is possible if and only if $\langle u_1, u_2\rangle = 0$. This means that u_1 and u_2 are orthogonal. Lastly, we need to prove that any symmetric matrix is diagonalizable. This is a little bit technical so we skip this proof.

□

Note that, in general, the eigenvectors of a matrix A are not orthogonal relative to the standard scalar product in $\mathbb{R}^n$ (or $\mathbb{C}^n$). However, if the matrix is symmetric then the eigenvectors are necessarily orthogonal[2], and they can be normalized, if needed. This yields a matrix of eigenvectors

$$P = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \quad \text{satisfying} \quad PP^T = I_n.$$

The condition $PP^T = I_n$ follows directly from $\langle u_i, u_j\rangle = \delta_{ij}$ (orthonormal eigenvectors). Hence the matrix P that contains the eigenvectors of a symmetric matrix is an orthogonal matrix.

Theorem 6. Let A be any $n \times n$ matrix. Then,

1. $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$,
2. $\text{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$,

where $\{\lambda_1, \ldots, \lambda_n\}$ are the eigenvalues of A counted with their multiplicity.

[2] Eigenvectors corresponding to different eigenvalues are necessarily orthogonal, while eigenvectors corresponding to the same eigenvalue with geometric multiplicity larger than one can be orthogonalized, e.g., by using Gram-Schmidt procedure.

Proof. To prove these identities, let us assume that A is diagonalizable[3]. In this case, we know that there exists a matrix P that has the eigenvectors of A as columns such that

$$A = P\Lambda P^{-1}, \qquad \text{where} \qquad \Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \tag{79}$$

is the diagonal matrix of eigenvalues. To prove 1. we simply notice that

$$\det(A) = \det(P\Lambda P^{-1}) = \det(P)\det(P^{-1})\det(\Lambda) = \det(\Lambda) = \lambda_1\lambda_2\cdots\lambda_n. \tag{80}$$

To prove 2. we notice that[4]

$$\text{trace}(A) = \text{trace}(P\Lambda P^{-1}) = \text{Tr}(PP^{-1}\Lambda) = \text{trace}(\Lambda) = \lambda_1 + \cdots + \lambda_n. \tag{81}$$

□

[3]The proof for the non-diagonalizable case is very much the same. The only difference is that we use the Jordan canonical form of A instead of the diagonal matrix of eigenvalues Λ.

[4]Recall that if A and B are two square matrices of the same size we have

$$\text{trace}(AB) = \text{trace}(BA).$$