Lecture 13: Some Properties of Eigenvalues and Eigenvector

We will continue the discussion on properties of eigenvalues and eigenvectors from Lecture 12. Throughout the present lecture, $A$ denotes an $n \times n$ matrix with real entries. A vector $v$, different from the zero-vector, is said to be an eigenvector if there is a scalar $\lambda$, such that

$$ Av = \lambda v. \quad (1) $$

The scalar $\lambda$ is referred to as an eigenvalue of $A$. The eigenvalue may be a complex number and the eigenvector may have complex entries. If $v$ is an eigenvector of $A$, then so is any nonzero multiple of this vector. In this lecture $I$ will denote the $n \times n$ identity matrix.

The eigenvalues are the zeros of the characteristic polynomial

$$ p_A(z) = \det(zI - A) \quad (2) $$

of $A$. This can be seen as follows: Let $\lambda$ be an eigenvalue of $A$. Then there is a nonvanishing vector $v$ that satisfies (1). The latter equation can be written in the form

$$ (A - \lambda I)v = 0. \quad (3) $$

The existence of a nonvanishing solution $v$ to this homogeneous equation shows that the matrix $A - \lambda I$ is singular. The determinant of a singular matrix vanishes. Therefore, $p_A(\lambda) = \det(A - \lambda I) = 0$. Conversely, assume that $p_A(\lambda) = 0$. Then the matrix $A - \lambda I$ is singular and the homogeneous equation (3) has a nonvanishing solution, which we denote by $v$. This vector satisfies (1). Hence, $\lambda$ is an eigenvalue and $v$ an associated eigenvector.

The fact that the characteristic polynomial (2) indeed is a polynomial can be seen by expanding the determinant along rows or columns. This shows that $p_A(z)$ is a polynomial of degree $n$. By the Fundamental Theorem of Algebra, a polynomial of degree $n$ has precisely $n$ zeros in the complex plane, counting multiplicities. Therefore an $n \times n$ matrix has precisely $n$ eigenvalues in the complex plane, counting multiplicities.

Example 1. Consider the matrix

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

Its characteristic polynomial is given by

$$ p_A(z) = (z - 1)(z - 1)(z - 2). $$

The eigenvalue $\lambda = 1$ is said to be of algebraic multiplicity 2, because it is a zero of $p_A(z)$ of multiplicity 2. The eigenvalue $\lambda = 2$ is of multiplicity 1. \quad \Box

Example 2. Expanding the characteristic polynomial for the matrix

$$ A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix} \quad (4) $$

along the last row yields

$$ p_A(z) = \det \begin{bmatrix} z - 1 & -2 \\ -1 & z - 3 \end{bmatrix} (z - 1) = (z^2 - 4z + 1)(z - 1). $$

\footnote{If $A - \lambda I$ were nonsingular, then the homogeneous equation (3) would have the unique solution $v = 0$.}
This shows that the eigenvalues of the matrix are
\[ \lambda_{1,2} = 2 \pm \sqrt{3}, \quad \lambda_3 = 1. \]

If we replace the entry 2 in position (1, 2) of the matrix (4) by any real number strictly smaller than −1, then the matrix has one pair of complex conjugate eigenvalues. For instance, setting the (1, 2)-entry to −2 yields the eigenvalues
\[ \lambda_{1,2} = 2 \pm i, \quad \lambda_3 = 1, \]
where \( i = \sqrt{-1} \) is the imaginary unit. □

Let \( S \) be a nonsingular \( n \times n \) matrix. Then the matrices \( A \) and \( SAS^{-1} \) are said to be similar. Similar matrices have the same eigenvalues. This can be seen as follows. Let
\[ p_{SAS^{-1}}(z) = \det(zI - SAS^{-1}) \]
de note the characteristic polynomial for the matrix \( SAS^{-1} \). Then
\[ p_{SAS^{-1}}(z) = \det(zSS^{-1} - SAS^{-1}) = \det(S(zI - A)S^{-1}) = \det(S)\det(zI - A)\det(S^{-1}). \quad (5) \]
Since
\[ \det(S)\det(S^{-1}) = \det(SS^{-1}) = \det(I) = 1, \]
the right-hand side of (5) simplifies to
\[ p_{SAS^{-1}}(z) = \det(zI - A), \]
which is the characteristic polynomial \( p_A(z) \) for the matrix \( A \). This shows that the matrices \( SAS^{-1} \) and \( A \) have the same characteristic polynomial and therefore the same eigenvalues.

The powerful QR-algorithm for computing all eigenvalues of a matrix \( A \), to be described in Lecture 14, carries out a sequence similarity transformations that yield a matrix \( SAS^{-1} \) of upper triangular form. The eigenvalues of an upper triangular matrix are given by the diagonal entries. Since \( SAS^{-1} \) and \( A \) are similar, the diagonal entries of \( SAS^{-1} \) are the eigenvalues of \( A \).

The geometric multiplicity of an eigenvalue \( \lambda \) is the number of linearly independent solutions of the homogeneous equation (3). Any eigenvalue has geometric multiplicity at least one, because every eigenvalue requires an eigenvector. However, the geometric multiplicity can be smaller than the algebraic multiplicity.

Example 3. The characteristic polynomial for the matrix
\[ A = \begin{bmatrix} -0.1 & 1 & 0 \\ 0 & -0.1 & 1 \\ 0 & 0 & -0.1 \end{bmatrix} \]
(6)
is given by
\[ p_A(z) = (z + 0.1)^3. \]
This shows that the matrix has the eigenvalue \( \lambda = -0.1 \) of algebraic multiplicity 3.

We seek to determine eigenvectors \( v = [\nu_1, \nu_2, \nu_3]^T \) associated with this eigenvalue by computing nontrivial solutions of the homogeneous linear system (3) with \( \lambda = -0.1 \). This system is of the form
\[ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (7) \]
The first equation of this system shows that \( \nu_2 = 0 \) and the second equation that \( \nu_3 = 0 \). The entry \( \nu_1 \) is arbitrary. We may, for instance, choose the eigenvector \( v = [1, 0, 0]^T \). There are no solutions of (7) that are not multiples of \( v \). Hence, there is only one linearly independent solution. Therefore the geometric multiplicity of the eigenvalue \( \lambda = -0.1 \) is only one. □

Matrices of order \( n \) that have fewer than \( n \) linearly independent eigenvectors are said to be \textit{defective}.

Example 4. If we change the \((3,1)\)-entry in the matrix (6) of Example 3 slightly, then the eigenvalues of the matrix obtained are simple, and every eigenvalue has an associated eigenvector. For instance, let

\[
A = \begin{bmatrix}
-0.1 & 1 & 0 \\
0 & -0.1 & 1 \\
1 \cdot 10^{-4} & 0 & -0.1 \\
\end{bmatrix}.
\]

(8)

Expanding the determinant that defines the characteristic polynomial \( p_A(z) \) for this matrix along its first column yields

\[
p_A(z) = (z + 0.1)^3 - 1 \cdot 10^{-4}.
\]

Its zeros can be computed explicitly. Let \( y = z + 0.1 \) and solve

\[
y^3 = 1 \cdot 10^{-4}.
\]

This equation has the solutions

\[
y_j = 1 \cdot 10^{-4/3} \exp(2\pi i (j - 1)/3), \quad j = 1, 2, 3,
\]

where \( i = \sqrt{-1} \). Therefore \( p_A(z) \) has the zeros

\[
z_j := y_j - 0.1 = -0.1 + 1 \cdot 10^{-4/3} \exp(2\pi i (j - 1)/3), \quad j = 1, 2, 3.
\]

Each eigenvalue has an associated eigenvector. Hence, the geometric multiplicity of each eigenvalue is one. □

It is quite easy to see that eigenvectors associated with distinct eigenvalues are linearly independent; see Exercise 1. Let the \( n \times n \) matrix \( A \) have eigenvalues \( \lambda_j \) and associated linearly independent eigenvectors \( v_j \), i.e.,

\[
A v_j = \lambda_j v_j, \quad 1 \leq j \leq n.
\]

(9)

The set of eigenvalues \( \{\lambda_j\}_{j=1}^n \) of \( A \) is referred to as the \textit{spectrum} of \( A \) and denoted by \( \lambda(A) \).

Introduce the diagonal matrix

\[
\Lambda = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \in \mathbb{R}^{n \times n}
\]

and the eigenvector matrix \( V = [v_1, v_2, \ldots, v_n] \in \mathbb{R}^{n \times n} \). Equations (9) can be written compactly in the form

\[
AV = V \Lambda.
\]

(10)

Since the columns of \( V \) are linearly independent, the inverse \( V^{-1} \) exists. Multiplying equation (10) by \( V^{-1} \) from the right-hand side yields the \textit{spectral decomposition} of \( A \),

\[
A = V \Lambda V^{-1}
\]

(11)

The spectral decomposition is useful for evaluating matrix functions. This has already been illustrated in Lecture 11, page 2.
Example 5. Let $A = \text{diag}[2,1,1]$. Then the axis vectors $e_j$, $1 \leq j \leq 3$, are eigenvectors and this gives the eigenvector matrix $V = I$. This eigenvector matrix is well-conditioned; it has condition number $1$. □

Example 6. Let $v_j$, $1 \leq j \leq 3$, be linearly independent eigenvectors of the matrix (8) of Example 4 and define the eigenvector matrix $V = [v_1, v_2, v_3]$. Since the matrix (8) is close to the defective matrix (6) of Example 3, its eigenvectors are nearly linearly dependent. The condition number of the eigenvector matrix, $\kappa_2(V) := \|V\|_2\|V^{-1}\|_2$, is $4.6 \cdot 10^2$. □

Defective matrices are fairly rare. There are many more matrices with nearly linearly dependent eigenvectors than defective ones.

Lecture 11 discussed the role of the eigenvalues of the matrix $A \in \mathbb{R}^{n \times n}$ for the stability of initial value problems for systems of ordinary differential equations

$$y'(t) = Ay(t), \quad t \geq 0, \quad y(0) = y_0,$$

where $'$ denotes the time-derivative $d/dt$ and $y_0 \in \mathbb{R}^n$ specifies the solution at the initial time $t = 0$. We showed that the eigenvalues have to have negative real part in order for the solution to decay as $t$ increases.

The following example illustrates that when $A$ is the defective matrix (6), two solution components first increase considerably when $t$ increases before starting to decrease, despite the fact that all eigenvalues are $-0.1$.

Example 7. Let the matrix $A$ in (12) be given by (6) and let $y_0 = [1,1,1]^T$. Figure 1 shows the 3 components of the solution vector

$$y(t) = \exp(At)y_0, \quad 0 \leq t \leq 2000.$$
The figure shows one of the solution components to grow from 1 to over $10^3$ before decaying towards zero as $t$ increases, despite the fact that all the eigenvalues of $A$ have negative real part. We remark that matrices $A$ with close eigenvalues and a well-conditioned eigenvector matrix will give solutions to the system of differential equations (12) that decay with increasing $t$ already for small values of $t$; see Exercise 2.

Systems of ordinary differential equations arise when modeling phenomena in, e.g., chemistry, biology, physics, and economy. Models only approximate reality. We therefore would like the solution to be insensitive to small perturbations in the model, i.e., in the matrix $A$ and initial values $y_0$. If the solution is very sensitive to perturbations in the matrix or initial values, then we cannot conclude much about the behavior of the system modeled.

![Figure 2: Components of the solution $y(t)$ of Example 7 as a function of $t$. Note the logarithmic scale on the vertical axis.](image)

Figure 2: Components of the solution $y(t)$ of Example 7 as a function of $t$. Note the logarithmic scale on the vertical axis.

Example 8. Let $A$ be the matrix (8) obtained the tiny perturbation $1 \cdot 10^{-4}$ of the $(3,1)$-entry of the matrix (6) used in Example 7. Figure 2 displays the solutions of the system of differential equations (12) determined by the matrix (8) with initial value $y_0 = [1, 1, 1]^T$. The solution can be seen to increase with $t$. This depends on that the tiny perturbation $10^{-4}$ of the matrix (6) moved one of the eigenvalues from $-0.01$ into the right complex half-plane. The system associated with the matrix (8) is not stable. This example illustrates that before drawing conclusions from mathematical model (12) one should investigate the sensitivity of the eigenvalues of the matrix. 

The set
$$\Lambda_\epsilon(A) := \{\lambda \in \lambda(A + E) : \text{for all } E \in \mathbb{R}^{n \times n} \text{ with } \|E\| \leq \epsilon\}$$

is referred to as the $\epsilon$-pseudospectrum of the matrix $A$. In order to be certain that the real system modeled
by a system of ordinary differential equations is stable, we should verify that the \( \epsilon \)-pseudospectrum of the matrix is in the left half of the complex plane for reasonable positive values of \( \epsilon \). The choice of \( \epsilon \) should depend on the accuracy of the model.

![Figure 3: The \( \epsilon \)-pseudospectrum of the matrix \( A \) of Example 3 with \( \epsilon = 1 \cdot 10^{-4} \) contains the red dots. The eigenvalues of \( A \) are marked by a circle.](image)

Example 9. In order to obtain an idea of the location of the \( \epsilon \)-pseudospectrum for \( \epsilon = 1 \cdot 10^{-4} \) of the matrix \( A \) of Example 3, we generate 1000 matrices \( E \in \mathbb{R}^{3 \times 3} \) with the MATLAB command \texttt{randn(3)} and normalize them to have norm \( \| E \|_2 = 1 \cdot 10^{-4} \). The eigenvalues of all the 1000 matrices \( A + E \) so obtained are depicted by (red) dots in Figure 3. The (black) circle shows the eigenvalues \( \lambda = -0.1 \) of the matrix \( A \). The figure indicates that the eigenvalues of \( A \) are very sensitive to perturbations in the matrix entries. The pattern of the dots depends on that the matrices \( A + E \) are real. \( \Box \)

There is an easy way to determine sets in the complex plane that contain all eigenvalues of a matrix \( A = [a_{jk}]_{j,k=1}^n \). Consider the \( n \) disks in the complex plane,

\[
D_j = \{ z : |z - a_{jj}| \leq \sum_{k \neq j} |a_{jk}| \}, \quad 1 \leq j \leq n.
\]

The sets \( D_j \) are known as \textit{Gershgorin disks}. All eigenvalue of the matrix \( A \) lie in the union of these disks. This bound for the eigenvalues is particularly useful when the diagonal entries of \( A \) are much larger in magnitude than the sum of the magnitude of the off-diagonal entries. Applications are illustrated in the exercises.

We note that if the radii of the Gershgorin disks are small, then the eigenvalues are not overly sensitive to small perturbations in the matrix entries. This follows from the observation that small perturbations of
the matrix entries give small perturbations of the radii of the Gershgorin disks. Hence, if the disks are small, then so are the disks for the perturbed matrix. The centers of the disks changes at most a little. Therefore the eigenvalues cannot move a lot when the matrix is perturbed.

We turn to real symmetric matrices, i.e., matrices $A \in \mathbb{R}^{n \times n}$ that satisfy $A = A^T$. The eigenvalues of a symmetric matrix are not very sensitive to perturbations of the matrix entries. The reason for this is that an $n \times n$ symmetric matrix has $n$ orthogonal eigenvectors. We will show that the eigenvectors of a symmetric matrix associated with distinct eigenvalues are orthogonal. Let $v_1$ and $v_2$ be eigenvectors of the symmetric matrix $A$ associated with the eigenvalues $\lambda_1$ and $\lambda_2$, and assume that

$$\lambda_1 \neq \lambda_2.$$  

(13)

Multiplying $Ax_1 = \lambda_1 x_1$ from the left by $x_2^T$ yields

$$\lambda_1 x_2^T x_1 = x_2^T (\lambda_1 x_1) = x_2^T (Ax_1) = x_2^T A x_1 = x_2^T A^T x_1 = (A x_2)^T x_1 = (\lambda_2 x_2)^T x_1 = \lambda_2 x_2^T x_1.$$

It now follows from (13) that $x_2^T x_1 = 0$, i.e., $x_1$ and $x_2$ are orthogonal. The vectors can be scaled to be of unit length. Then they are orthonormal.

Eigenvectors associated with the same eigenvalue are not uniquely determined, but can be chosen to be orthonormal. This gives an orthonormal basis of eigenvectors. Letting these orthonormal eigenvectors form the columns of the matrix $V$ gives an orthogonal matrix. Orthogonal matrices have condition number 1; see Exercise 3.

Example 10. The $n \times n$ identity matrix is symmetric and has the eigenvalue 1 of algebraic multiplicity 1. Any nonvanishing vector is an eigenvector. We therefore can choose the axis vectors $e_j$ as eigenvectors. They form an orthonormal basis of $\mathbb{R}^n$. This shows in particular that the geometric multiplicity of the eigenvalue 1 is $n$. 

Example 11. The matrix $A = \text{diag}[2, 1, 1]$ is symmetric. The vector $e_1 = [1, 0, 0]^T$ is an eigenvector associated with the eigenvalue 2. Any nonvanishing vector of the form $v = [0, \nu_2, \nu_3]^T$ is an eigenvector associated with the eigenvalue 1. We can choose the orthonormal vectors $e_2 = [0, 1, 0]^T$ and $e_3 = [0, 0, 1]^T$ as eigenvectors. There are also other orthonormal eigenvectors, as well as linearly independent nonorthogonal eigenvectors associated with the eigenvalue 1. 

Symmetric matrices only have real eigenvalues. This can be seen as follows. Assume that $\lambda$ is a complex eigenvalue and $v$ an associated eigenvector, which necessarily has complex entries. Then

$$Av = \lambda v, \quad A\bar{v} = \bar{\lambda} \bar{v},$$

where the bar denotes complex conjugation. Then, in view of that $A^T = A$, we obtain

$$\lambda \bar{v}^T v = \bar{v}^T (\lambda v) = \bar{v}^T (Av) = (\bar{v}^T A) v = (\bar{v}^T A^T) v = (A \bar{v})^T v = (\bar{\lambda} \bar{v})^T v = \bar{\lambda} \bar{v}^T v.$$

Since $v$ is not the zero-vector, the expression $\bar{v}^T v$ is nonvanishing. It follows that $\lambda = \bar{\lambda}$. This only can hold if the imaginary part of $\lambda$ vanishes, i.e., if $\lambda$ is real.

**Exercises**

Exercise 1. Let $A$ be a $2 \times 2$ matrix with distinct eigenvalues. Show that the eigenvectors associated with the different eigenvalues are linearly dependent. Hint: Assume that the eigenvectors are linearly dependent and seek to obtain a contradiction. Bonus proof: Show that an $n \times n$ matrix with $n$ distinct eigenvalues has $n$ linearly independent eigenvectors. 

7
Exercice 2. Let $\Lambda = \text{diag}[-0.11, -0.1, -0.09]$ and determine a random orthogonal $3 \times 3$ matrix $V$. This can be done with the MATLAB command $V = \text{orth(randn}(3))$. Form the matrix $A = V\Lambda V^T$. This matrix has the eigenvalues $\{ -0.11, -0.1, -0.09 \}$. Why? Plot the solution of the system of ordinary differential equations (12) defined by this matrix and initial values $y_0 = [1, 1, 1]^T$. How does the behavior of the solutions differ from those shown in Examples 7 and 8.

Exercice 3. Show that an orthogonal matrix $V$ has condition number 1. Hint: Show that $\|V\|_2 = 1$ and $\|V^T\|_2 = 1$ by using the definition of a matrix norm.

Exercice 4. What are the Gershgorin disks for the matrix

$$A = \begin{bmatrix} 1 & 0.1 & -0.1 \\ 0.01 & 0.5 & -0.01 \\ 0 & 0.01 & 0.1 \end{bmatrix}.$$

Is the matrix singular? Justify your answer.

Exercice 5. What are the Gershgorin disks for the matrices (4) and (6).

Exercice 6. Plot the $\epsilon$-pseudospectrum for $\epsilon = 0.01$ and $\epsilon = 0.001$ for the tridiagonal $100 \times 100$ matrix with zeros on the diagonal, 1 on the superdiagonal, and 0.25 on the subdiagonal. Also plot the eigenvalues. Compute the condition number of the eigenvector matrix.

Exercice 7. Plot the $\epsilon$-pseudospectrum of the for $\epsilon = 0.01$ for the matrix $\text{diag}[0.1, 0.2, 0.3, \ldots, 10]$. What is the condition number of the eigenvector matrix?