PART V - Eigenvalues

Lecture 24 - *Eigenvalue Problems*

OBJECTIVE:
Eigenvalue problems form an important class of problems in scientific computing.
The algorithms to solve them are powerful, yet far from obvious!
Here we review the theory of eigenvalues and eigenvectors.
Algorithms are discussed in later lectures.

◇ EIGENVALUES AND EIGENVECTORS

Let $\mathbf{A} \in \mathbb{R}^{m \times m}$, $\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^m$.

Then $\mathbf{x}$ is an *eigenvector* of $\mathbf{A}$ and $\lambda \in \mathbb{R}$ its corresponding *eigenvalue* if

$$\mathbf{Ax} = \lambda \mathbf{x}$$

The idea is that the action of $\mathbf{A}$ on a subspace $\mathbf{S}$ of $\mathbb{R}^m$ can act like scalar multiplication.

This special subspace $\mathbf{S}$ is called an *eigenspace*.

The set of all the eigenvalues of a matrix $\mathbf{A}$ is called the *spectrum* of $\mathbf{A}$, denoted $\Lambda(\mathbf{A})$.

Broadly speaking, eigenvalues and eigenvectors are useful for two reasons:

1. Algorithmically, eigenvalue analysis can decouple a set of equations to a collection of scalar ones.

2. Physically, eigenvalue analysis can give insight into the evolution of linear systems of differential equation, in particular the stability of equilibria.

◇ EIGENVALUE DECOMPOSITION

We have already mentioned this in Lecture 5.

An eigenvalue decomposition of $\mathbf{A} \in \mathbb{R}^{m\times m}$ is a factorization

$$\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$$

where $\mathbf{X}$ is nonsingular and $\mathbf{\Lambda}$ is diagonal.

Note 1. *Such a decomposition does not always exist!*

The definition can be rewritten as

$$\mathbf{A}\mathbf{X} = \mathbf{X}\boldsymbol{\Lambda}$$

i.e.,

$$\begin{bmatrix} \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \mid \mathbf{x}_2 \mid \ldots \mid \mathbf{x}_n \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{x}_1 \mid \mathbf{x}_2 \mid \ldots \mid \mathbf{x}_m \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}$$

This makes it clear that

$$\mathbf{A}\mathbf{x}_j = \lambda_j \mathbf{x}_j$$

i.e., the j^{th} column of $\mathbf{X}$ is the j^{th} eigenvector and the (j, j) entry of $\boldsymbol{\Lambda}$ is the corresponding eigenvalue.

The eigenvalue decomposition expresses a change of basis to "eigenvector coordinates".

i.e., if $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{A} = \mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1}$, then we can write

$$(\mathbf{X}^{-1}\mathbf{b}) = \boldsymbol{\Lambda}(\mathbf{X}^{-1}\mathbf{x})$$

So to compute $\mathbf{A}\mathbf{x}$, we can expand $\mathbf{x}$ in the basis of columns of $\mathbf{X}$, apply $\boldsymbol{\Lambda}$, and interpret the results as coefficients of a linear combination of the columns of $\mathbf{X}$.

$\diamond$ GEOMETRIC MULTIPLICITY

The *geometric multiplicity* of an eigenvalue λ is the number of linearly independent eigenvectors associated with it.

The set of eigenvectors corresponding to a single eigenvalue (plus the zero vector) forms a subspace of $\mathbb{R}^m$ known as an *eigenspace*.

If $\lambda \in \boldsymbol{\Lambda}(\mathbf{A})$, let us denote the corresponding eigenspace by $\mathbf{E}_\lambda$.

$\mathbf{E}_\lambda$ is an *invariant subspace of* $\mathbf{A}$
i.e.,

$$\mathbf{A}\mathbf{E}_\lambda \subseteq \mathbf{E}_\lambda$$

The dimension of $\mathbf{E}_\lambda$ can then be interpreted as geometric multiplicity of λ.

i.e., the maximum number of linearly independent eigenvectors that can be found for a given λ.

◇ CHARACTERISTIC POLYNOMIAL

The *characteristic polynomial* $p_{\mathbf{A}}$ of $\mathbf{A} \in \mathbb{R}^{m \times m}$ is the degree-m polynomial

$$p_{\mathbf{A}}(z) = \det(z\mathbf{I} - \mathbf{A})$$

Note 2. *This polynomial is* monic*; i.e., the coefficient of z^m is 1.*

Theorem 1. *λ is an eigenvalue of $\mathbf{A}$ if and only if $p_{\mathbf{A}}(\lambda) = 0$.*

Note 3. Even if $\mathbf{A}$ is real, λ could be complex! However, if $\mathbf{A}$ is real, any complex λ must appear in complex conjugate pairs.
i.e., if $\mathbf{A}$ is real and $\lambda = a + ib$ is an eigenvalue, then so is $\lambda^* = a - ib$.

$\diamond$ ALGEBRAIC MULTIPLICITY

Since $p_{\mathbf{A}}(z)$ is a monic degree-m polynomial, it can be written as

$$p_{\mathbf{A}}(z) = (z - \lambda_1)(z - \lambda_2) \ldots (z - \lambda_m)$$

for some numbers $\lambda_j \in \mathbb{C}$ (the roots of the polynomial).

Each λ_j is an eigenvalue of $\mathbf{A}$, and in general may be repeated.
e.g.,

$$\lambda^2 - 2\lambda + 1 = (\lambda - 1)(\lambda - 1)$$

We define the *algebraic multiplicity* of an eigenvalue λ as the multiplicity of λ as a root of $p_{\mathbf{A}}(z)$.

An eigenvalue is *simple* if its algebraic multiplicity is 1.

Theorem 2. *If* $\mathbf{A} \in \mathbb{R}^{m\times m}$*, then* $\mathbf{A}$ *has* m *eigenvalues counting algebraic multiplicity.*
In particular, if the roots of $p_{\mathbf{A}}(z)$ *are simple, then* $\mathbf{A}$ *has* m *distinct eigenvalues.*

Note 4. *Every matrix has at least one eigenvalue! (why?)*

We will see shortly that the algebraic multiplicity of an eigenvalue is always at least as large as its geometric multiplicity (this may already be obvious!).

◇ SIMILARITY TRANSFORMATIONS

If $\mathbf{X} \in \mathbb{R}^{m\times m}$ is nonsingular, then

$$\mathbf{A} \longrightarrow \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

is called a *similarity transformation* of $\mathbf{A}$.

We say that two matrices $\mathbf{A}$ and $\mathbf{B}$ are *similar* if there is a similarity transformation of one to the other.

i.e., there is a nonsingular $\mathbf{X} \in \mathbb{R}^{m\times m}$ such that

$$\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

Many properties are shared by matrices that are similar.

Theorem 3. *If $\mathbf{X}$ is nonsingular, then $\mathbf{A}$ and $\mathbf{X}^{-1}\mathbf{A}\mathbf{X}$ have the same characteristic polynomial, eigenvalues, and algebraic and geometric multiplicities.*

Theorem 4. *The algebraic multiplicity of an eigenvalue λ is at least as large as its geometric multiplicity.*

◇ DEFECTIVE EIGENVALUES AND MATRICES

A generic matrix will have algebraic and geometric multiplicities that are equal (to 1) since eigenvalues are often not repeated.

However, this is certainly not true of every matrix!

Consider,

$$\mathbf{A} = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}$$

Both $\mathbf{A}$ and $\mathbf{B}$ have a single eigenvalue $\lambda = 2$ with algebraic multiplicity 3. *(verify!)*

For $\mathbf{A}$, we can choose 3 linearly independent eigenvectors (try $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$). So, the geometric multiplicity of $\mathbf{A}$ is 3.

However, for $\mathbf{B}$, we only have 1 linearly independent eigenvector (try $\mathbf{e}_1$). So, the geometric multiplicity of $\mathbf{B}$ is 1.

An eigenvalue whose algebraic multiplicity is greater than its geometric multiplicity is called a *defective eigenvalue*.

A matrix that has at least one defective eigenvalue is a *defective matrix*;
i.e., it does not possess a full set of m linearly independent eigenvectors.

Every diagonal matrix is non-defective, with algebraic multiplicity of every eigenvalue λ equal to its geometric multiplicity (equal to the number of times it occurs on the diagonal).

◇ DIAGONALIZABILITY

Non-defective matrices are precisely those matrices that have an eigenvalue decomposition.

Theorem 5. $\mathbf{A} \in \mathbb{R}^{m \times m}$ *is non-defective if and only if it has an eigenvalue decomposition*

$$\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$$

In view of this, another term for non-defective is *diagonalizable*.

◇ DETERMINANT AND TRACE

Both the *trace* of $\mathbf{A} \in \mathbb{R}^{m\times m}$ ($\mathrm{tr}(\mathbf{A}) = \sum_{j=1}^{m} a_{jj}$) and its determinant are related simply to its eigenvalues.

Theorem 6.

$$\mathrm{tr}(\mathbf{A}) = \sum_{j=1}^{m} \lambda_j \qquad \textit{and} \qquad \det(\mathbf{A}) = \prod_{j=1}^{m} \lambda_j$$

◇ ORTHOGONAL DIAGONALIZATION

It is possible not only for $\mathbf{A} \in \mathbb{R}^{m\times m}$ to have m linearly independent eigenvectors, but these vectors can be chosen to be orthogonal.

In such a case we say $\mathbf{A}$ is *orthogonally diagonalizable*; i.e., there exists an orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$$

Note 5. *Such a decomposition is both an eigenvalue decomposition and a SVD (except possibly for the signs of the elements of $\mathbf{\Lambda}$).*

Theorem 7. *A symmetric matrix is orthogonally diagonalizable and its eigenvalues are real.*

This is not the only class of orthogonally diagonalizable matrices.

It turns out that the entire class of orthogonally diagonalizable matrices has an elegant characterization.

We say that a matrix is *normal* if $\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T$. Then we have . . .

Theorem 8. *A matrix is orthogonally diagonalizable if and only if it is normal.*

$\diamond$ SCHUR FACTORIZATION

This final factorization is actually the most useful in numerical analysis because all matrices (even defective ones) have a *Schur factorization*

$$\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^T$$

where $\mathbf{Q}$ is orthogonal and $\mathbf{T}$ is upper-triangular.

Note 6. *Since* $\mathbf{A}$ *and* $\mathbf{T}$ *are similar, the eigenvalues of* $\mathbf{A}$ *appear on the diagonal of* $\mathbf{T}$ (why?)

Theorem 9. *Every square matrix* $\mathbf{A}$ *has a Schur factorization.*

$\diamond$ EIGENVALUE-REVEALING FACTORIZATIONS

We have just described three *eigenvalue-revealing factorizations*;
i.e., factorizations where the matrix is reduced to a form where the eigenvalues can simply be read off.

We summarize them as follows:

1. A diagonalization $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ exists if and only if $\mathbf{A}$ is non-defective.

2. An orthogonal diagonalization $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ exists if and only if $\mathbf{A}$ is normal.

3. An orthogonal triangularization (Schur decomposition) $\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^T$ always exists.

To compute eigenvalues, we will construct one of these factorizations.

In general, we will use the Schur decomposition since it applies to all matrices without restriction and it uses orthogonal transformations, which have good stability properties.

If $\mathbf{A}$ is normal, then its Schur factorization will have a diagonal $\mathbf{T}$.

Moreover, if $\mathbf{A}$ is symmetric, we can exploit this symmetry to reduce $\mathbf{A}$ to diagonal form with half as much work or less than is required for general $\mathbf{A}$.