Mathematics, Pusan National University

Numerical Linear Algebra

Lecture 24. Eigenvalue Problems

Taehyeong Kim
th_kim@pusan.ac.kr

October 26, 2020

Content

Definition 1.1

Let $A \in \mathbb{C}^{m\times m}$ be a square matrix. A nonzero vector $x \in \mathbb{C}^m$ is an *eigenvector* of A, and $\lambda \in \mathbb{C}$ is its corresponding *eigenvalue*, if

$$Ax = \lambda x.$$

The set of all the eiganvalues of matrix A is the *spectrum* of A denoted by $\Lambda(A)$.

Eigenvalue Decomposition

Definition 2.1

An *eigenvalue decomposition* of a square matrix A, already mentioned in section 5, is a factorization

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

This definition can be rewritten

$$AX = X\Lambda,$$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}\begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}.$$

The jth column of X is an eigenvector of A and the j th entry of A is the corresponding eigenvalue.

$$Ax_j = \lambda_j x_j$$

Definition 3.1

The *characteristic polynomial* of $A \in \mathbb{C}^{m\times m}$, denoted by p_A of simply p, is the degree m polynomial defined by

$$p_A(z) = \det(zI - A).$$

Additionly, p is monic, i.e. the cofficient of its degree m term is 1.

Folling theorem has an important consequence. Even if a matrix is real, some of its eigenvalues may be complex. Physically, this is related to the phenomenon that real dynarnicd systems can have motions that oscillate as well as grow or decay. Algorithmically, it means that even if the input to a matrix eigenvalue problem is real, the output may have to be complex.

Characteristic Polynomial

Theorem 3.2

λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$

Proof.

This follows from the definition of an eigenvalue:

$$
\begin{aligned}
\lambda \text{ is an eigenvalue } &\Leftrightarrow \text{ there is a nonzero vector } x \text{ such that } \lambda x - Ax = 0 \\
&\Leftrightarrow \lambda I - A \text{ is singular} \\
&\Leftrightarrow \det(\lambda I - A) = 0
\end{aligned}
$$

□

Algebraic Multiplicity

By fundamental theorem of algebra, we can write p_A in the form

$$p_A(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)$$

for some numbers $\lambda_j \in \mathbb{C}$.
We define the *algebraic multiplicity* of an eigenvalue λ of A to be its multiplicity as a root of p_A. An eigenvalue is *simple* if its algebraic multiplicity is 1.

Remark

A number a is a root of a polynomial P if and only if the linear polynomial $x - a$ devides P, that is another polynomial Q such that $P = (x - a)Q$.

$$\begin{cases} a \text{ is called a } \textit{multiple} \text{ root of } P : (x-a)^2 \text{ devides } P \\ a \text{ is called a } \textit{simple} \text{ root of } P : \text{otherwise} \end{cases}$$

Algebraic Multiplicity

Theorem 4.1

If $A \in \mathbb{C}^{m\times m}$, then A has m eigenvalues, counted with algebric multiplicity. Inparticular, if the roots of p_A are simple, then A has m distinct eigenvalues.

Note

Every matrix has at least one eigenvalue.

Geometric Multiplicity

If λ is an eigenvalue of A, let us denote the corresponding eigenspace by E_λ. An eigenspace E_λ is an example of an *invariant subspace* of A; that is $AE_\lambda \subseteq E_\lambda$.
The dimension of E_λ can be interpreted as the maximum number of linear independent eigenvectors that can be found, all with the same eigenvalue λ. The number is known as the *geometric multiplicity* of λ. The geometric multiplicity can also ve described as the dimension of the nullspace of $A - \lambda I$, since that nullspace is again E_λ.
Suppose λ is an eigenvalue of A.

1. **Eigenvectors (geometric)** There are nonzero solutions to $A\mathbf{x} = \lambda\mathbf{x}$.
2. **Eigenvalues (algebraic)** The determinant of $A - \lambda I$ is zero.

And we want to know its **multiplicity**.

1. **(Geometric Multiplicity = GM)** Count the independent eigenvectors for λ. Look at the dimension of the nullspace of $A - \lambda I$.
2. **(Algebraic Multiplicity = AM)** Count the repetitions of λ among the eigenvalues. Look at the roots of $det(A - \lambda I) = 0$.

Similarity Transformations

If $X \in \mathbb{C}^{m \times m}$ is nonsingular, then the map $A \to X^{-1}AX$ is called a similarity transformation of A.

Definition 6.1

We say two matrices A and B are *similar* if there is a similarity transformation relating one to the other, i.e.

$$\text{there exists a nonsingular } X \in \mathbb{C}^{m \times m} \text{ such that } B = X^{-1}AX.$$

There are many properties about similar matrices A and $X^{-1}AX$.

Similarity Transformations

Theorem 6.2

If X is nonsingular, then A and $X^{-1}AX$ have the same characteristic polynomial, eigenvalues, and algebraic and geometric multiplicies.

Proof.

The characteristic polynomials match is a straightforward computation:

$$\begin{aligned} p_{X^{-1}AX}(z) &= \det(zI - X^{-1}AX) = \det(X^{-1}(zI - A)X) \\ &= \det(X^{-1})\det(zI - A)\det(X) = \det(zI - A) = p_A(z). \end{aligned}$$

□

We can now relate geometric multiplicity to algebraic multiplicity.

Theorem 6.3

The algebruic multiplicity of an eigenvalue λ is at least as great as its geometric multiplicity.

Proof.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_r$ be linearly independent eigenvectors associated to $\hat{\lambda}$, so $\hat{\lambda}$ has geometric multiplicity r. Let $\mathbf{x}_{r+1}, \ldots, \mathbf{x}_n$ be basis for $\mathbb{R}^n$.
And let X be the matrix which columns $\mathbf{x}_k$.
Consider AX.

$$AX = \begin{bmatrix} | & & | & \\ \hat{\lambda}\mathbf{x}_1 & \cdots & \hat{\lambda}\mathbf{x}_r & \cdots \\ | & & | & \end{bmatrix} \Rightarrow S^{-1}AS = \begin{bmatrix} \hat{\lambda}I & B \\ 0 & C \end{bmatrix}$$

where $B : r \times n$ matrix, $C : (n-r) \times (n-r)$ matrix.

Proof.

By theorem 6.2, the characteristic polynomial of A and $S^{-1}AS$ are the same.
It is easy to see that the characteristic polynomial of $S^{-1}AS$ has a factor of at least $(\hat{\lambda} - \lambda)^r$.($\because$ determinant of block matrices)
$\therefore$ GM $\leq$ AM. $\square$

Example 7.1

Consider the matrices

$$A = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$

Both A and B have same characteristic polynomial $(z-3)^3$, so there is a single eigenvalue $\lambda = 2$ of algebraic multiplicity 3. In the case of A, we can choose three independent eigenvectors, e_1, e_2, e_3. So the geometric multiplicity is also 3. But for B, we can find only a single independent eigenvector, so the geometric multiplicity of the eigenvalue is only 1.

An eigenvalue whose algebraic multiplicity exceeds its geometric multiplicity is a defective eigenvalue. A matrix that has one or more defective eigenvalues is a defective matrix.

Diagonalisability

Theorem 8.1

An $m \times m$ matrix A is nondefective if and only if it has an eigenvalue decomposition $A = X^{-1}\Lambda X$.

Proof.

($\Leftarrow$) Given an eigenvalue decomposition $A = X^{-1}\Lambda X$, Λ is similar to A. Since Λ is a diagonal matrix, it is nondefective, and thus the same holds for A.
($\Rightarrow$) A nondefective matrix must have m linearly independent eigenvectors, because eigenvectors with different eigenvalues must be linear independent, and each eigenvalue cam contribute as many linearly independent eigenvectors as its multiplicity. If these m independent eigenvectors are formed into the columns of a matrix X, then X is nonsingular and we have $A = X^{-1}\Lambda X$. □

In view of this result, another term for nondefective is daagonalzzable.

Question

Does a diagonalizable matrix A in some sense "behave like" its diagonal equivalent Λ?

Answer

The answer depends on what aspect of behavior one measures and on the condition number of X, the matrix of eigenvectors.

If X is highly ill-conditioned, then a great deal of information may be discarded in passing from A to Λ.("A Note of Caution: Nonnormality" in Lecture 34.)

Definition 9.1

The *trace* of $A \in \mathbb{C}^{m\times m}$ is the sum of its diagonal elements:

$$\mathrm{tr}(A) = \sum_{j=1}^{m} a_{jj}.$$

Both the trace and the determinant are related simply to the eigenvalues.

Determinant and Trace

Theorem 9.2

The determinant $\det(A)$ *and the trace* $tr(A)$ *are equar to the product and the sum of the eigenvalues of* A*, respectively, counted with algebraic multiplicity:*

$$\det(A) = \prod_{j=1}^{m} \lambda_j, \quad tr(A) = \sum_{j=1}^{m} \lambda_j.$$

Proof.

$$\det(A) = (-1)^m \det(-A) = (-1)^m p_A(0) = \prod_{j=1}^{m} \lambda_j.$$

By definition of characteristic polynomial, the cofficient of the z^{m-1} term of p_A is the negative of the sum of the diagonal elements of A, or $-\mathrm{tr}(A)$.

And also $p_A(z) = (z - \lambda_1)\cdots(z - \lambda_m)$, so $\mathrm{tr}(A) = \sum_{j=1}^{m} \lambda_j$. □

Unitary Diagonalization

If we can choose m orthogonal eigenvector fo $m \times m$ matrix A, A is *unitarily diagonalizable*:

there exists a unitary matrix Q such that $A = Q\Lambda Q^*$

Theorem 10.1

A hermitian matrix is unitarily diagonalizable, and its eigenvalues are real.

Theorem 10.2

A matrix is unitarily diagonalizable, if and only if it is normal.

Unitary Diagonalization

There is some equivalent definition of normal matrix.

1. A is normal.
2. A is diagonalizable by a unitary matrix.
3. There exists a set of eigenvector of A which forms an orthonormal basis for $\mathbb{C}^m$.
4. $\|Ax\| = \|A^*x\|$ for every x.
5. The Frobenius norm of A can be computed by the eigenvalues of A : $\operatorname{tr}(A^*A) = \sum_j |\lambda_j|^2$.
6. The Hermitian part $1/2(A + A^*)$ and skew-Hermitian part $1/2(A - A^*)$ of A commute.
7. $A^* = AU$ for some unitary matrix U.
8. U and P comute, where we have the polar decomposition $A = UP$ with a unitary matrix U and some positive semidefinite matrix P.
9. $\sigma_j = |\lambda_j|$ for all j.

Schur Factorization

Definition 11.1

A *Schur factorization* of a matrix A is a factorization

$$A = QTQ^*,$$

where Q is unitary and T is upper-triangular.

Note

A and T are similar.

Schur Factorization

Theorem 11.2

Every square matrix A has a Schur factorization.

Proof.

We will show that there exists triangular matrix T such that $T = UTU^*$ for some unitary marix U for all A.
We use mathematical induction on size of A.
$(n = 1)$ trivial.
Assume that $n > 1$, and the result holds for all matrices of size less than n. n. Since every complex matrix has an eigenvalue, choose an eigenvalue λ of A and an associated eigenvector $\mathbf{v} = (v_1, \ldots, v_n)$.
Let $\mathbf{x} = \frac{\overline{v_1}\mathbf{v}}{\|\overline{v_1}\mathbf{v}\|}$, and set $u = \mathbf{x} - e_1$.

Proof.

And we will put Q in some cases.

$$\begin{cases} Q : \text{Householder matrix associated with } u & (\text{if } \mathbf{x} \neq e_1) \\ Q = I & (\text{if } \mathbf{x} = e_1) \end{cases}$$

Then $\mathbf{x} = Qe_1$, it means that the first column of Q is $\mathbf{x}$. We already know that every householder matrix is unitary and hermitian. So x^* is first row of Q^*. Since $Q^{-1} = Q^* = Q$, $Q = \left[\mathbf{x}|V\right] = \begin{bmatrix} \mathbf{x}^* \\ V^* \end{bmatrix}$. Therefore,

$$QAQ = QA\left[\mathbf{x}|V\right] = Q\left[\lambda\mathbf{x}|AV\right] = \left[\lambda e_1 \middle| \begin{bmatrix} \mathbf{x}^* \\ V^* \end{bmatrix} AV\right] = \begin{bmatrix} \lambda & \mathbf{x}^* AV \\ \mathbf{0} & V^* AV \end{bmatrix}.$$

Proof.

The size of V^*AV is $(n-1)\times(n-1)$, so we can apply the induction, there exists unitary matrix R such that $T_{n-1} = R^*(V^*AV)R$ is upper triangular matrix. Let

$$U = Q\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix},$$

then

$$U^*U = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} Q^*Q \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} = I.$$

So U is unitary.

Schur Factorization

Proof.

$$
\begin{aligned}
T = U^*AU &= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} QAQ \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} \\
&= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} \begin{bmatrix} \lambda & \mathbf{x}^*AV \\ \mathbf{0} & V^*AV \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} \\
&= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ \mathbf{0} & V^*AVR \end{bmatrix} \\
&= \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ \mathbf{0} & R^*V^*AVR \end{bmatrix} = \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ \mathbf{0} & T_{n-1} \end{bmatrix}
\end{aligned}
$$

Hence, T is triangular matrix.

$\therefore A = UTU^*$ □

Eigenvalue-Revealing Factorizations

We cam summarize this Lecture as follows.

- A diagonalization $A = X\Lambda X^{-1}$ exists if and only if A is nondefective.
- A unitary diagonalization $A = Q\Lambda Q^*$ exists if and only if A is normal.
- A unitary triangularization (Schur factorization) $A = QTQ^*$ always exists.

To compute eigenvalues, we shall construct one of these factorizations. In general, this will be the Schur factorization, since this applies without restriction to all matrices. Moreover, since unitary transformations are involved, the algorithms that result tend to be numerically stable. If A is normal, then the Schur form comes out diagonal, and in particular, if A is hermitian, then we can take advantage of this symmetry throughout the computation and reduce A to diagonal form with half as much work or less than is required for general A.

Thank you!