CS724: Topics in Algorithms
Eigenvalues of Matrices

Prof. Dan A. Simovici

Definition

Let $A \in \mathbb{C}^{m \times m}$ be a square matrix. A pair $(\lambda, \mathbf{x})$ that consists of a complex number $\lambda \in \mathbb{C}$ and a non-zero complex vector $\mathbf{x}$ is an *eigenpair* if $A\mathbf{x} = \lambda\mathbf{x}$. The number λ is an *eigenvalue* of A, and $\mathbf{x}$ is an *eigenvector* of A.
The set of eigenvalues of A is known as the *spectrum* of A and is denoted by $\text{spec}(A)$. The *spectral radius* of A is the number

$$\rho(A) = \max\{|\lambda| \mid \lambda \in \text{spec}(A)\}.$$

Theorem

Let $A \in \mathbb{C}^{n \times n}$ *be a matrix and let* $\lambda_1, \ldots, \lambda_n$ *distinct eigenvalues of* A. *If* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *are eigenvectors of* A *that that correspond to distinct eigenvalues, then the set* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *is linearly independent.*

Proof

If the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ were not linearly independent we would have a linear combination of these vectors

$$c_{i_1}\mathbf{v}_{i_1} + \cdots + c_{i_p}\mathbf{v}_{i_p} = \mathbf{0}_n,$$

containing a minimal number of vectors such that not every one of scalars $c_{i_1}, \ldots, c_{i_p}$ is 0. This implies

$$\begin{aligned} & c_{i_1}A\mathbf{v}_{i_1} + \cdots + c_{i_p}A\mathbf{v}_{i_p} \\ & \quad = \quad c_{i_1}\lambda_{i_1}\mathbf{v}_{i_1} + \cdots + c_{i_p}\lambda_{i_p}\mathbf{v}_{i_p} = \mathbf{0}. \end{aligned}$$

These equalities imply

$$c_{i_1}(\lambda_{i_1} - \lambda_{i_p})\mathbf{v}_{i_1} + \cdots + c_{i_{p-1}}(\lambda_{i_{p-1}} - \lambda_{i_p})\mathbf{v}_{i_{p-1}} = \mathbf{0}_n,$$

which contradicts the minimality of the number of terms. Thus, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.

Corollary

If $A \in \mathbb{C}^{n \times n}$ then the set of eigenvalues of A does not contain more than n distinct eigenvalues.

Proof.

Since the maximum size of a linearly independent set in $\mathbb{C}^n$ is n, it follows that A cannot have more than n distinct eigenvalues. □

Definition

Let $A \in \mathbb{C}^{n \times n}$ be a matrix. A subspace S of $\mathbb{C}^n$ is a *right invariant subspace of the matrix A* if $A\mathbf{s} \in S$ for every $\mathbf{s} \in S$, and is a *left invariant subspace* if $A^{\mathsf{H}}\mathbf{s} \in S$ for every $\mathbf{s} \in S$.

Definition

The *geometric multiplicity of an eigenvalue* λ of a matrix $A \in \mathbb{C}^{n\times n}$ is denoted by $\text{geomm}(A, \lambda)$ and is equal to $\dim(S_{A,\lambda})$.
Equivalently, the geometric multiplicity of λ is

$$\text{geomm}(A, \lambda) = \dim(\text{NullSp}(A - \lambda I_n)) = n - rank(A - \lambda I_n).$$

Theorem

Let $A \in \mathbb{R}^{n\times n}$*. We have* $0 \in spec(A)$ *if and only if* A *is a singular matrix. Moreover, in this case,* $geomm(A, 0) = n - rank(A) = \dim(NullSp(A))$*.*

Example

The matrix I_n has 1 as its unique eigenvalue. Its invariant subspace is the entire space V; therefore, the geometric multiplicity of 1 is $\dim(V)$.

Definition

Let $A \in \mathbb{C}^{m \times n}$ be a matrix. An *invariant subspace of A* is a subspace S of $\mathbb{R}^n$ such that $\mathbf{x} \in S$ implies $A\mathbf{x} \in S$.

- the null space of a matrix A is an invariant subspace;
- if $\mathbf{x}$ is an eigenvector of A, then $\{a\mathbf{x} \mid a \in \mathbb{C}\}$ is an invariant subspace of A.

If λ is an eigenvalue of $A \in \mathbb{C}^{n \times n}$ we have $\mathbf{x}^{\mathsf{H}} A \mathbf{x} = \lambda \mathbf{x}^{\mathbf{H}} \mathbf{x}$, so

$$\lambda = \frac{\mathbf{x}^{\mathsf{H}} A \mathbf{x}}{\mathbf{x}^{\mathsf{H}} \mathbf{x}}.$$

In the real case we replace $\mathbf{x}^{\mathsf{H}}$ by $\mathbf{x}'$: if $A \in \mathbb{R}^{n \times n}$, λ is an eigenvalue and $\mathbf{x}$ is an eigenvector that corresponds to λ, then

$$\lambda = \frac{\mathbf{x}' A \mathbf{x}}{\mathbf{x}' \mathbf{x}}.$$

Theorem

Let $A \in \mathbb{C}^{n \times n}$ and let $S \subseteq \mathbb{C}^n$ be an invariant subspace of A. If the columns of a matrix $X \in \mathbb{C}^{n \times p}$ constitute a basis of S, then there exists a unique matrix $L \in \mathbb{C}^{p \times p}$ such that $AX = XL$.

Proof

Let $X = (\mathbf{x}_1 \ \cdots \ \mathbf{x}_p)$. Since $A\mathbf{x}_1 \in S$ it follows that $A\mathbf{x}_1$ can be uniquely expressed as a linear combination of the columns of X, that is,

$$A\mathbf{x}_j = \mathbf{x}_1\ell_{1j} + \cdots + \mathbf{x}_p\ell_{pj}$$

for $1 \leqslant i \leqslant p$. Thus,

$$A\mathbf{x}_j = X \begin{pmatrix} \ell_{1j} \\ \vdots \\ \ell_{pj} \end{pmatrix}.$$

The matrix L is defined by $L = (\ell_{ij})$.

Corollary

$(\lambda, \mathbf{v})$ *is an eigenpair of* L *if and only if* $(\lambda, X\mathbf{v})$ *is an eigenpair of* A.

UMASS BOSTON

Let $A \in \mathbb{C}^{n\times n}$ be a matrix having the eigenvalues $\lambda_1, \ldots, \lambda_n$. If $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are n eigenvectors corresponding to these values, then we have

$$A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1, \ldots, A\mathbf{x}_n = \lambda_n \mathbf{x}_n.$$

By introducing the matrix $X = (\mathbf{x}_1 \ \cdots \ \mathbf{x}_n) \in \mathbb{C}^{n\times n}$ these equalities can be written in a concentrated form as

$$AX = X\text{diag}(\lambda_1, \ldots, \lambda_n).$$

Obviously, since the eigenvalues can be listed in several ways, this equality is not unique.

Suppose now that $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are unit vectors and that the eigenvalues $\lambda_1, \ldots, \lambda_n$ are distinct. Then X is a unitary matrix, $X^{-1} = X^{\mathsf{H}}$ and we obtain the equality

$$A = X \mathrm{diag}(\lambda_1, \ldots, \lambda_n) X^{\mathsf{H}} = \lambda_1 \mathbf{x}_1 \mathbf{x}_1^{\mathsf{H}} + \cdots + \lambda_n \mathbf{x}_n \mathbf{x}_n^{\mathsf{H}}$$

known as the *spectral decomposition* of the matrix A.

If λ is an eigenvalue of the matrix $A \in \mathbb{C}^{n\times n}$, there exists a non-zero eigenvector $\mathbf{x} \in \mathbb{C}^n$ such that $A\mathbf{x} = \lambda\mathbf{x}$. Therefore, the linear system

$$(\lambda I_n - A)\mathbf{x} = \mathbf{0}_n$$

has a non-trivial solution. This is possible if and only if $\det(\lambda I_n - A) = 0$, so eigenvalues are the solutions of the equation

$$\det(\lambda I_n - A) = 0.$$

Note that $\det(\lambda I_n - A)$ is a polynomial of degree n in λ, known as the *characteristic polynomial* of the matrix A. We denote this polynomial by p_A.

Example

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

be a matrix in $\mathbb{C}^{3\times 3}$. Its characteristic polynomial is

$$\begin{aligned} p_A &= \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix} = \lambda^3 - (a_{11} + a_{22} + a_{33})\lambda^2 \\ & +(a_{11}a_{22} + a_{22}a_{33} + a_{33}a_{11} - a_{12}a_{21} - a_{23}a_{32} - a_{13}a_{31})\lambda \\ & -(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11} - a_{13}a \end{aligned}$$

Theorem

Let $A \in \mathbb{C}^{n\times n}$. Then, $spec(A) = spec(A')$ and $spec(A^H) = \{\bar{\lambda} \mid \lambda \in spec(A)\}$.

Proof: We have

$$p_{A'}(\lambda) = \det(\lambda I_n - A') = \det((\lambda I_n - A)') = \det(\lambda I_n - A) = p_A(\lambda).$$

Thus, since A and A' have the same characteristic polynomials, their spectra are the same.
For A^{H} we can write

$$p_{A^{\mathrm{H}}}(\bar{\lambda}) = \det(\bar{\lambda} I_n - A^{\mathrm{H}}) = \det((\lambda I_n - A)^{\mathrm{H}}) = (p_A(\lambda))^{\mathrm{H}},$$

which implies the second part of the Theorem.

The characteristic polynomial of a matrix can be computed in **R** using the function `charpoly` of the `pracma` package.
For the matrix A defined in

```
> A <- matrix(c(1:6),3,3)
> A
     [,1] [,2] [,3]
[1,]    1    4    1
[2,]    2    5    2
[3,]    3    6    3
```

the characteristic polynomial is $\lambda^3 - 9\lambda^2$ as returned by

```
> charpoly(A)
[1]  1 -9  0  0
```

Let B be the matrix defined as

```
> B <- matrix(c(1,0,2,3,1,1,1,4,2),3,3)
> B
     [,1] [,2] [,3]
[1,]    1    3    1
[2,]    0    1    4
[3,]    2    1    2
```

If `charpoly` is called as in

```
> charpoly(B,info=TRUE)
```

then, in addition to the characteristic polynomial of B, its determinant and inverse matrix are also returned as in

```
$cp
[1]   1  -4  -1 -20

$det
[1] 20

$inv
     [,1]  [,2]  [,3]
[1,] -0.1 -0.25  0.55
[2,]  0.4  0.00 -0.20
[3,] -0.1  0.25  0.05
```

To compute the eigenvalues of a matrix one could use the `eigen` function of the base package of **R** .
The following call to `eigen` computes the eigenvalues of the matrix A together with its characteristic vectors:

```
> A <- matrix(c(1:6),3,3)
> eigen(A)
eigen() decomposition

$values
[1]  9.000000e+00  2.497182e-09 -2.497182e-09

$vectors
          [,1]          [,2]          [,3]
[1,] 0.3713907 -7.071068e-01  7.071068e-01
[2,] 0.5570860 -1.177183e-09 -1.177183e-09
[3,] 0.7427814  7.071068e-01 -7.071068e-01
```

Equality of spectra of A and A' does not imply that the eigenvectors or the invariant subspaces of the corresponding eigenvalues are identical, as it can be seen from the following example.

Example

Consider the matrix $A \in \mathbb{C}^{2\times 2}$ defined by

$$A = \begin{pmatrix} a & 0 \\ c & b \end{pmatrix},$$

where $a \neq b$ and $c \neq 0$. It is immediate that $\text{spec}(A) = \text{spec}(A') = \{a, b\}$. For $\lambda_1 = a$ we have the distinct invariant subspaces:

$$\begin{aligned} S_{A,a} &= \left\{ k \begin{pmatrix} a-b \\ c \end{pmatrix} \middle| k \in \mathbb{C} \right\} \\ S_{A',a} &= \left\{ k \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| k \in \mathbb{C} \right\}, \end{aligned}$$

as the reader can easily verify.

- The leading term of the characteristic polynomial of A is generated by $(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$ and equals λ^n.
- The fundamental theorem of algebra implies that p_A has n complex roots, not necessarily distinct. Observe also that, if A is a matrix with real entries, the roots are paired as conjugate complex numbers.

Definition

The *algebraic multiplicity of an eigenvalue* λ of a matrix $A \in \mathbb{C}^{n \times n}$, $\text{algm}(A, \lambda)$ equals k if λ is a root of order k of the equation $p_A(\lambda) = 0$. If $\text{algm}(A, \lambda) = 1$, we refer to λ as a *simple eigenvalue*.

Let $A \in \mathbb{R}^{3\times 3}$ be the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix}$$

The characteristic polynomial of A is

$$p_A(\lambda) = \begin{vmatrix} \lambda - 1 & -1 & -1 \\ 0 & \lambda - 1 & -2 \\ -2 & -1 & \lambda \end{vmatrix} = \lambda^3 - 2\lambda^2 - 3\lambda.$$

Therefore, the eigenvalues of A are $3, 0$ and -1.

The eigenvalues of I_3 are obtained from the equation

$$det(\lambda I_3 - I_3) = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^3 = 0.$$

Thus, I_3 has one eigenvalue, 1, and $\text{algm}(I_3, 1) = 3$.

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a matrix and let $\lambda \in spec(A)$. Then, for any $k \in \mathbb{P}$, $\lambda^k \in spec(A^k)$.

Proof.

The proof is by induction on $k \geqslant 1$. The base step, $k = 1$ is immediate. Suppose that $\lambda^k \in \text{spec}(A^k)$, that is $A^k\mathbf{x} = \lambda^k\mathbf{x}$ for some $\mathbf{x} \in V - \{\mathbf{0}\}$. Then, $A^{k+1}\mathbf{x} = A(A^k\mathbf{x}) = A(\lambda^k\mathbf{x}) = \lambda^k A\mathbf{x} = \lambda^{k+1}\mathbf{x}$, so $\lambda^{k+1} \in \text{spec}(A^{k+1})$. □

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a non-singular matrix and let $\lambda \in spec(A)$. We have $\frac{1}{\lambda} \in spec(A^{-1})$ and the sets of eigenvectors of A and A^{-1} are equal.

Proof.

Since $\lambda \in \text{spec}(A)$ and A is non-singular we have $\lambda \neq 0$ and $A\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \in V - \{\mathbf{0}\}$. Therefore, we have $A^{-1}(A\mathbf{x}) = \lambda A^{-1}\mathbf{x}$, which is equivalent to $\lambda^{-1}\mathbf{x} = A^{-1}\mathbf{x}$, which implies $\frac{1}{\lambda} \in \text{spec}(A^{-1})$. In addition, this implies that the set of eigenvectors of A and A^{-1} are identical. □

UMASS BOSTON

Theorem

Let $p_A(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$ *be the characteristic polynomial of the matrix A. Then, we have* $c_i = (-1)^i S_i(A)$ *for* $1 \leqslant i \leqslant n$*, where* $S_i(A)$ *is the sum of all principal minors of order i of A.*

Proof

Since $p_A(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$, it is easy to see that the derivatives of $p_A(\lambda)$ are given by:

$$\begin{aligned}
p_A^{(1)}(\lambda) &= n\lambda^{n-1} + (n-1)c_1\lambda^{n-2} + \cdots + c_{n-1},\\
p_A^{(2)}(\lambda) &= n(n-1)\lambda^{n-2} + (n-1)(n-2)c_1\lambda^{n-3} + \cdots + 2c_{n-2}),\\
&\vdots\\
p_A^{(k)}(\lambda) &= n(n-1)\cdots(n-k+1)\lambda^{n-k} + \cdots + k!c_{n-k}),\\
&\vdots\\
p_A^{(n)}(\lambda) &= n!c_0.
\end{aligned}$$

This implies

$$c_{n-k} = k!p_A^{(k)}(0)$$

for $0 \leqslant k \leqslant n$.
On other hand, $c_{n-k} = \frac{1}{k!}(-1)^k k! S_{n-k}(A) = (-1)^{n-k} S_{n-k}(A)$, which implies the statement of theorem.

By Viéte's Theorem, taking into account Theorem 17 we have:

$$\lambda_1 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn} = trace(A) = -c_1.$$

Another interesting fact is

$$\lambda_1 \cdots \lambda_n = \det(A).$$

Theorem

Let $A \in \mathbb{C}^{m \times n}$ *and* $B \in \mathbb{C}^{n \times m}$ *be two matrices. Then the set of non-zero eigenvalues of the matrices* $AB \in \mathbb{C}^{m \times m}$ *and* $BA \in \mathbb{C}^{n \times n}$ *are the same and* $algm(AB, \lambda) = algm(BA, \lambda)$ *for each such eigenvalue.*

Proof

Consider the following straightforward equalities:

$$\begin{pmatrix} I_m & -A \\ O_{n,m} & \lambda I_n \end{pmatrix} \begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix} = \begin{pmatrix} \lambda I_m - AB & O_{m,n} \\ -\lambda B & \lambda I_n \end{pmatrix}$$

$$\begin{pmatrix} -I_m & O_{m,n} \\ -B & \lambda I_n \end{pmatrix} \begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix} = \begin{pmatrix} -\lambda I_m & -A \\ O_{n,m} & \lambda I_n - BA \end{pmatrix}.$$

Observe that

$$\det\left(\begin{pmatrix} I_m & -A \\ O_{n,m} & \lambda I_n \end{pmatrix} \begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix}\right) = \det\left(\begin{pmatrix} -I_m & O_{m,n} \\ -B & \lambda I_n \end{pmatrix} \begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix}\right),$$

and therefore,

$$\det \begin{pmatrix} \lambda I_m - AB & O_{m,n} \\ -\lambda B & \lambda I_n \end{pmatrix} = \det \begin{pmatrix} -\lambda I_m & -A \\ O_{n,m} & \lambda I_n - BA \end{pmatrix}.$$

The last equality amounts to

$$\lambda^n p_{AB}(\lambda) = \lambda^m p_{BA}(\lambda).$$

Thus, for $\lambda \neq 0$ we have $p_{AB}(\lambda) = p_{BA}(\lambda)$, which gives the desired conclusion.

Corollary

Let

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

be a vector in $\mathbb{C}^n - \{\mathbf{0}\}$. *Then, the matrix* $\mathbf{a}\mathbf{a}^H \in \mathbb{C}^{n \times n}$ *has one eigenvalue distinct from* 0, *and this eigenvalue is equal to* $\| \mathbf{a} \|^2$.

Theorem

Let $A \in \mathbb{C}^{(m+n)\times(m+n)}$ *be a matrix partitioned as*

$$A = \begin{pmatrix} B & C \\ O_{n,m} & D \end{pmatrix},$$

where $B \in \mathbb{C}^{m\times m}$, $C \in \mathbb{C}^{m\times n}$, *and* $D \in \mathbb{C}^{n\times n}$. *Then,* $spec(A) = spec(B) \cup spec(D)$.

Proof

Let $\lambda \in \text{spec}(A)$ and let $\mathbf{x} \in \mathbb{C}^{m+n}$ be an eigenvector that corresponds to λ. If

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix},$$

where $\mathbf{u} \in \mathbb{C}^m$ and $\mathbf{v} \in \mathbb{C}^n$, then we have

$$A\mathbf{x} = \begin{pmatrix} B & C \\ O_{n,m} & D \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} B\mathbf{u} + C\mathbf{v} \\ D\mathbf{v} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}.$$

This implies $B\mathbf{u} + C\mathbf{v} = \lambda\mathbf{u}$ and $D\mathbf{v} = \lambda\mathbf{v}$. If $\mathbf{v} \neq \mathbf{0}$, then $\lambda \in \text{spec}(D)$; otherwise, $B\mathbf{u} = \lambda\mathbf{u}$, which yields $\lambda \in \text{spec}(B)$, so $\lambda \in \text{spec}(B) \cup \text{spec}(D)$. Thus, $\text{spec}(A) \subseteq \text{spec}(B) \cup \text{spec}(D)$.

To prove the converse inclusion, note that if $\lambda \in \text{spec}(B)$ and $\mathbf{u}$ is an eigenvector of λ, then $B\mathbf{u} = \lambda\mathbf{u}$, which means that

$$A \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix},$$

so $\text{spec}(B) \subseteq \text{spec}(A)$. Similarly, $\text{spec}(D) \subseteq \text{spec}(A)$, which implies the equality of the theorem.

Theorem

All eigenvalues of a Hermitian matrix $A \in \mathbb{C}^{n \times n}$ are real numbers.
All eigenvalues of a skew-Hermitian matrix are purely imaginary numbers.

Proof.

Note that $\mathbf{x}^{\mathsf{H}}\mathbf{x}$ is a real number for every $\mathbf{x} \in \mathbb{C}^n$. Since $\lambda = \frac{\mathbf{x}^{\mathsf{H}} A\mathbf{x}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}}$, λ is a real number.

Suppose now that B is a skew-Hermitian matrix. Then, as above, $\overline{\mathbf{x}^{\mathsf{H}} A\mathbf{x}} = -\mathbf{x}^{\mathsf{H}} A\mathbf{x}$, which implies that the real part of $\mathbf{x}^{\mathsf{H}} A\mathbf{x}$ is 0. Thus, $\mathbf{x}^{\mathsf{H}} A\mathbf{x}$ is a purely imaginary number and λ is a purely imaginary number. □

Corollary

If $A \in \mathbb{R}^{n\times n}$ and A is a symmetric matrix, then all its eigenvalues are real numbers.

Proof.

This statement follows from Theorem 21 by observing that the Hermitian adjoint A^{H} of a matrix $A \in \mathbb{R}^{n\times n}$ coincides with its transposed matrix A'. □

Corollary

Let $A \in \mathbb{C}^{m\times n}$ be a matrix. The non-zero eigenvalues of the matrices AA^H and A^HA are positive numbers and they have the same algebraic multiplicities for the matrices AA^H and A^HA.

Proof

We need to verify only that if λ is a non-zero eigenvalue of $A^{\mathsf{H}}A$, then λ is a positive number. Since $A^{\mathsf{H}}A$ is a Hermitian matrix, λ is a real number. The equality $A^{\mathsf{H}}A\mathbf{x} = \lambda\mathbf{x}$ for some eigenvector $\mathbf{x} \neq \mathbf{0}$ implies

$$\lambda \parallel \mathbf{x} \parallel_2^2 = \lambda \mathbf{x}^{\mathsf{H}}\mathbf{x} = (A\mathbf{x})^{\mathsf{H}} A\mathbf{x} = \parallel A\mathbf{x} \parallel_2^2,$$

so $\lambda > 0$.

Corollary

Let $A \in \mathbb{C}^{m \times n}$ be a matrix. The eigenvalues of the matrix $B = A^H A \in \mathbb{C}^{n \times n}$ are real non-negative numbers.

The matrix B defined above is clearly Hermitian and, therefore, its eigenvalues are real numbers. Next, if λ is an eigenvalue of B, then

$$\lambda = \frac{\mathbf{x}^{\mathsf{H}} A^{\mathsf{H}} A \mathbf{x}}{\mathbf{x}^{\mathsf{H}} \mathbf{x}} = \frac{(A\mathbf{x})^{\mathsf{H}} A \mathbf{x}}{\mathbf{x}^{\mathsf{H}} \mathbf{x}} = \frac{\| A\mathbf{x} \|}{\| \mathbf{x} \|} \geqslant 0,$$

where $\mathbf{x}$ is an eigenvector that corresponds to λ.

If A is a Hermitian matrix, then $A^{\mathsf{H}}A = A^2$, hence the spectrum of $A^{\mathsf{H}}A$ is $\{\lambda^2 \mid \lambda \in \text{spec}(A)\}$.

Theorem

If $A \in \mathbb{C}^{n\times n}$ is a Hermitian matrix and $\mathbf{u}, \mathbf{v}$ are two eigenvectors that correspond to two distinct eigenvalues λ_1 and λ_2, then $\mathbf{u} \perp \mathbf{v}$.

Proof.

We have $A\mathbf{u} = \lambda_1 \mathbf{u}$ and $A\mathbf{v} = \lambda_2 \mathbf{v}$. This allows us to write $\mathbf{v}^{\mathsf{H}} A\mathbf{u} = \lambda_1 \mathbf{v}^{\mathsf{H}} \mathbf{u}$. Since A is Hermitian, we have

$$\lambda_1 \mathbf{v}^{\mathsf{H}} \mathbf{u} = \mathbf{v}^{\mathsf{H}} A\mathbf{u} = \mathbf{v}^{\mathsf{H}} A^{\mathsf{H}} \mathbf{u} = (A\mathbf{v})^{\mathsf{H}} \mathbf{u} = \lambda_2 \mathbf{v}^{\mathsf{H}} \mathbf{u},$$

which implies $\mathbf{v}^{\mathsf{H}} \mathbf{u} = 0$, that is, $\mathbf{u} \perp \mathbf{v}$. □

Theorem

If $A, B \in \mathbb{C}^{n \times n}$ and $A \sim B$, then the two matrices have the same characteristic polynomials and, therefore, spec$(A) =$ spec(B).

Proof.

Since $A \sim B$, there exists an invertible matrix X such that $A = XBX^{-1}$. Then, the characteristic polynomial $\det(A - \lambda I_n)$ can be rewritten as

$$\begin{aligned} \det(A - \lambda I_n) &= \det(XBX^{-1} - \lambda X I_n X^{-1}) \\ &= \det\left(X(B - \lambda I_n)X^{-1}\right) \\ &= \det(X)\det(B - \lambda I_n)\det(X^{-1}) \\ &= \det(B - \lambda I_n), \end{aligned}$$

which implies $\text{spec}(A) = \text{spec}(B)$. □

Theorem

If $A, B \in \mathbb{C}^{n \times n}$ and $A \sim B$, then $trace(A) = trace(B)$.

Proof.

Since the two matrices are similar, they have the same characteristic polynomials, so both $trace(A)$ and $trace(B)$ equal $-c_1$, where c_1 is the coefficient of λ^{n-1} in both $p_A(\lambda)$ and $p_B(\lambda)$. □

Theorem

If $A \sim_u B$*, where* $A, B \in \mathbb{C}^{n\times n}$*, then the Frobenius norm of these matrices are equal, that is,* $\| A \|_F = \| B \|_F$.

Proof.

Since $A \sim_u B$, there exists a unitary matrix U such that $A = UBU^{\mathsf{H}}$. Therefore,

$$A^{\mathsf{H}}A = UB^{\mathsf{H}}U^{\mathsf{H}}UBU^{\mathsf{H}} = UB^{\mathsf{H}}BU^{\mathsf{H}},$$

which implies $A^{\mathsf{H}}A \sim_u B^{\mathsf{H}}B$. Therefore, these matrices have the same characteristic polynomials which allows us to infer that $trace(A^{\mathsf{H}}A) \sim_u trace(B^{\mathsf{H}}B)$, which yields the desired equality. □

UMASS BOSTON

Theorem

Let $A \in \mathbb{C}^{n\times n}$ *and* $B \in \mathbb{C}^{k\times k}$ *be two matrices. If there exists a matrix* $U \in \mathbb{C}^{n\times k}$ *having an orthonormal set of columns such that* $AU = UB$*, then there exists* $V \in \mathbb{C}^{n\times(n-k)}$ *such* $(U\ V) \in \mathbb{C}^{n\times n}$ *is a unitary matrix and*

$$(U\ V)^H A (U\ V) = \begin{pmatrix} B & U^H AV \\ O & V^H AV \end{pmatrix}.$$

Proof

Since U has an orthonormal set of columns, there exists $V \in \mathbb{C}^{n\times(n-k)}$ such that $(U\ V)$ is a unitary matrix.
We have

$$\begin{aligned} U^{\mathsf{H}}AU &= U^{\mathsf{H}}UB = I_kB = B, \\ V^{\mathsf{H}}AU &= V^{\mathsf{H}}UB = OB = O, \end{aligned}$$

which allows us to write

$$\begin{aligned} (U\ V)^{\mathsf{H}}A(U\ V) &= (U\ V)^{\mathsf{H}}(AU\ AV) = \begin{pmatrix} U^{\mathsf{H}} \\ V^{\mathsf{H}} \end{pmatrix}(AU\ AV) \\ &= \begin{pmatrix} U^{\mathsf{H}}AU & U^{\mathsf{H}}AV \\ V^{\mathsf{H}}AU & V^{\mathsf{H}}AV \end{pmatrix} = \begin{pmatrix} B & U^{\mathsf{H}}AV \\ O & V^{\mathsf{H}}AV \end{pmatrix}. \end{aligned}$$

Corollary

Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix and $B \in \mathbb{C}^{k\times k}$ be a matrix. If there exists a matrix $U \in \mathbb{C}^{n\times k}$ having an orthonormal set of columns such that $AU = UB$, then there exists $V \in \mathbb{C}^{n\times(n-k)}$ such that $(U\ V)$ is a unitary matrix and

$$(U\ V)^{\mathsf{H}} A (U\ V) = \begin{pmatrix} B & O \\ O & V^{\mathsf{H}} A V \end{pmatrix}.$$

Proof.

Since A is Hermitian we have $U^{\mathsf{H}} A V = U^{\mathsf{H}} A^{\mathsf{H}} V = (V^{\mathsf{H}} A U)^{\mathsf{H}} = O$, which produces the desired result. □

Corollary

Let $A \in \mathbb{C}^{n\times n}$, λ be an eigenvalue of A, and let $\mathbf{u}$ be an eigenvector of A with $\| \mathbf{u} \| = 1$ that corresponds to λ. There exists $V \in \mathbb{C}^{n\times(n-1)}$ such that $(\mathbf{u}\ V) \in \mathbb{C}^{n\times n}$ is a unitary matrix and

$$(\mathbf{u}\ V)^H A (\mathbf{u}\ V) = \begin{pmatrix} \lambda & \mathbf{u}^H A V \\ \mathbf{0}_{n-1} & V^H A V \end{pmatrix}.$$

If A is a Hermitian matrix, then

$$(\mathbf{u}\ V)^H A (\mathbf{u}\ V) = \begin{pmatrix} \lambda & \mathbf{0}_{n-1} \\ \mathbf{0}_{n-1} & V^H A V \end{pmatrix}.$$

Theorem

(Schur's Triangularization Theorem) *Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. There exists a unitary matrix $Q \in \mathbb{C}^{n \times n}$ and an upper-triangular matrix $T \in \mathbb{C}^{n \times n}$ such that $A = QTQ^H$ and the diagonal elements of T are the eigenvalues of A. Moreover, each eigenvalue λ occurs in the sequence of diagonal values a number of $\text{algm}(A, \lambda)$ times.*

Proof

The argument is by induction on $n \geqslant 1$. The base case, $n = 1$, is trivial. So, suppose that the statement is true for matrices in $\mathbb{C}^{(n-1)\times(n-1)}$. Let $\lambda_1 \in \mathbb{C}$ be an eigenvalue of A, and let $\mathbf{u}$ be an eigenvector that corresponds to this eigenvalue. We have

$$Q^{\mathsf{H}} A Q = \begin{pmatrix} \lambda_1 & \mathbf{u}^{\mathsf{H}} A V \\ \mathbf{0}_{n-1} & V^{\mathsf{H}} A V \end{pmatrix},$$

where $Q = (\mathbf{u}|V)$ is an unitary matrix.
By the inductive hypothesis, since $V^{\mathsf{H}} A V \in \mathbb{C}^{(n-1)\times(n-1)}$, there exists a unitary matrix $S \in \mathbb{C}^{(n-1)\times(n-1)}$ such that $V^{\mathsf{H}} A V = S^{H} W S$, where W is an upper-triangular matrix.

Proof (cont'd)

Then, we have

$$Q^{\mathsf{H}} A Q = \begin{pmatrix} \lambda_1 & \mathbf{u}^{\mathsf{H}} V S^{\mathsf{H}} W S \\ \mathbf{0}_{n-1} & S^{\mathsf{H}} W S \end{pmatrix} = \begin{pmatrix} \lambda_1 & O \\ \mathbf{0}_{n-1} & W \end{pmatrix},$$

which shows that an upper triangular matrix T that is unitarily similar to A can be defined as

$$T = \begin{pmatrix} \lambda_1 & O \\ \mathbf{0}_{n-1} & W \end{pmatrix}.$$

Proof (cont'd)

Since $T \sim_u A$, it follows that the two matrices have the same characteristic polynomials and therefore, the same spectra and algebraic multiplicities for each eigenvalue.

Example

Let $A \in \mathbb{R}^{3\times 3}$ be the symmetric matrix

$$A = \begin{pmatrix} 14 & -10 & -2 \\ -10 & -5 & 5 \\ -2 & 5 & 11 \end{pmatrix}$$

whose characteristic polynomial is:

$$p_A(\lambda) = \lambda^3 - 20\lambda^2 - 100\lambda + 2000.$$

The eigenvalues of A are $\lambda_1 = 20, \lambda_2 = 10$ and $\lambda_3 = -10$.
It is easy to see that

$$\mathbf{v}_1 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -2 \\ -5 \\ 1 \end{pmatrix}$$

are eigenvectors that correspond to the eigenvalues λ_1, λ_2 and λ_3, respectively.

Example

The corresponding unit vectors are

$$\mathbf{u}_1 = \begin{pmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} -\frac{2}{\sqrt{30}} \\ -\frac{5}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \end{pmatrix}.$$

For $Q = (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3)$ we have

$$Q'AQ = Q'(20\mathbf{u}_1 \ 10\mathbf{u}_2 \ -10\mathbf{u}_3) = \text{diag}(20, 10, -10).$$

The Schur decomposition of a square matrix can be computed in **R** using the function `Schur` of the package `Matrix`.
For the matrix A considered before, we can write:

```
> A <- matrix(c(14,-10,-2,-10,-5,5,-2,5,11),3,3)
> A
     [,1] [,2] [,3]
[1,]   14  -10   -2
[2,]  -10   -5    5
[3,]   -2    5   11
```

The call to the function `Schur`

```
> Schur(A,vectors=TRUE)
```

returns a result that has the following components:

```
$Q
           [,1]       [,2]          [,3]
[1,]  0.3651484  0.8164966  4.472136e-01
[2,]  0.9128709 -0.4082483 -3.750263e-19
[3,] -0.1825742 -0.4082483  8.944272e-01

$T
     [,1]          [,2]          [,3]
[1,]  -10 -1.831868e-15 -1.160892e-15
[2,]    0  2.000000e+01  7.604338e-16
[3,]    0  0.000000e+00  1.000000e+01

$EValues
[1] -10  20  10
```

If the `vectors` parameter is set to `FALSE` the result includes `$T` and `$EValues`.

Corollary

Let $A \in \mathbb{C}^{n \times n}$ *and let* f *be a polynomial. If* $spec(A) = \{\lambda_1, \ldots, \lambda_n\}$ *(including multiplicities), then* $spec(f(A)) = \{f(\lambda_1), \ldots, f(\lambda_n)\}$.

Proof

By Schur's Triangularization Theorem there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ and an upper-triangular matrix $T \in \mathbb{C}^{n \times n}$ such that $A = U^{\mathsf{H}} T U$ and the diagonal elements of T are the eigenvalues of A, $\lambda_1, \ldots, \lambda_n$. Therefore $Uf(A)U^{-1} = f(T)$, and the diagonal elements of $f(T)$ are $f(\lambda_1), , \ldots, f(\lambda_m)$. Since $f(A) \sim f(T)$, we obtain the desired conclusion because two similar matrices have the same eigenvalues with the same algebraic multiplicities.

The next statement presents a property of real matrices that admit real Schur factorizations.

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a real square matrix. If there exists a orthogonal matrix $U \in \mathbb{R}^{n\times n}$ and an upper-triangular matrix $T \in \mathbb{R}^{n\times n}$ such that $A = U^{-1}TU$, that is, a real Schur factorization, then the eigenvalues of A are real numbers.

Proof.

If the above factorization exists we have $T = UAU^{-1}$. Thus, the eigenvalues of A are the diagonal components of T and, therefore, they are real numbers. □