Math 6610: Analysis of Numerical Methods, I
Eigenvalues and eigenvectors

Department of Mathematics, University of Utah

Fall 2025

Resources: Trefethen and Bau 1997, Lecture 24
Süli and Mayers 2003, Section 5.1
Salgado and Wise 2022, Sections 1.3, 8.3

Eigenvalues and eigenvectors

Given $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $(\lambda, \boldsymbol{v}) \in \mathbb{C} \times (\mathbb{C}^n \backslash \{0\})$ is an eigenvalue-eigenvector pair if

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

Recall: it doesn't matter what value(s) λ takes on, but $\boldsymbol{v}$ *cannot* be $\mathbf{0}$.

Given $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $(\lambda, \boldsymbol{v}) \in \mathbb{C} \times (\mathbb{C}^n \backslash \{0\})$ is an eigenvalue-eigenvector pair if

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

Recall: it doesn't matter what value(s) λ takes on, but $\boldsymbol{v}$ *cannot* be $\boldsymbol{0}$.

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

Some additional terminology/properties:

- The collection of all eigenvalues of $\boldsymbol{A}$ is $\lambda(\boldsymbol{A}) \subset \mathbb{C}$, its *spectrum*.
- Even if $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\lambda(\boldsymbol{A})$ can contain complex values, and eigenvectors can be complex-valued.
- On paper, we typically identify the spectrum by computing roots of the *characteristic polynomial*, $p_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A})$. (This turns out to be a *terrible* algorithmic strategy.)

Eigenpair properties

There are many properties of eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{C}^{n \times n}$:

Eigenpair properties

There are many properties of eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{C}^{n \times n}$:

- All square matrices have exactly n eigenvalues, with some possibly repeated.
- All square matrices have at least 1 eigenvector.

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

(Eigenvectors aren't unique,
because of scaling)

Eigenpair properties

There are many properties of eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{C}^{n \times n}$:

- All square matrices have exactly n eigenvalues, with some possibly repeated.
- All square matrices have at least 1 eigenvector.
- Eigenvectors from *distinct* eigenvalues are linearly independent.
- Every *distinct* eigenvalue has at least 1 eigenvector.

why?

$A v_1 = \lambda_1 v_1$

$A v_2 = \lambda_2 v_2$ $\quad \lambda_1 \neq \lambda_2$

Consider $c_1 v_1 + c_2 v_2$ (arbitrary c_1, c_2)

$c_1 v_1 + c_2 v_2 \overset{?}{=} 0$

$$\left.\begin{aligned} c_1 \lambda_1 v_1 + c_2 \lambda_1 v_2 &\overset{?}{=} 0 \\ c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 &= 0 \end{aligned}\right\} \Longrightarrow c_1 = c_2 = 0$$

A

There are many properties of eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{C}^{n \times n}$:

- All square matrices have exactly n eigenvalues, with some possibly repeated.
- All square matrices have at least 1 eigenvector.
- Eigenvectors from *distinct* eigenvalues are linearly independent.
- Every *distinct* eigenvalue has at least 1 eigenvector.
- The collection of all eigenvectors associated to an eigenvalue λ is a subspace, and is frequently called an *eigenspace* E_λ.
- Eigenspaces are invariant subspaces of $\boldsymbol{A}$, i.e., $\boldsymbol{A}E_\lambda \subseteq E_\lambda$.

$$\underline{\underline{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad E_{\underline{1}} = \text{span}\{\underline{e}_1, \underline{e}_2\}$$

$E = \mathbb{C}^3$ is an invariant subspace of $\underline{\underline{A}}$

There are many properties of eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{C}^{n \times n}$:

- All square matrices have exactly n eigenvalues, with some possibly repeated.
- All square matrices have at least 1 eigenvector.
- Eigenvectors from *distinct* eigenvalues are linearly independent.
- Every *distinct* eigenvalue has at least 1 eigenvector.
- The collection of all eigenvectors associated to an eigenvalue λ is a subspace, and is frequently called an *eigenspace* E_λ.
- Eigenspaces are invariant subspaces of $\boldsymbol{A}$, i.e., $\boldsymbol{A}E_\lambda \subseteq E_\lambda$.
- The number of times an eigenvalue is repeated a_λ is its *algebraic multiplicity*
- The *geometric multiplicity* g_λ of an eigenvalue λ is $\dim E_\lambda$.
- *Simple* eigenvalues λ have $g_\lambda = a_\lambda = 1$.

$$\sum_{\lambda \in \lambda(A)} a_\lambda = n$$

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \lambda(A) = \{1\}$$

$$a_1 = 2, \quad g_1 = 1$$

Eigenpair properties

There are many properties of eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{C}^{n \times n}$:

- All square matrices have exactly n eigenvalues, with some possibly repeated.
- All square matrices have at least 1 eigenvector.
- Eigenvectors from *distinct* eigenvalues are linearly independent.
- Every *distinct* eigenvalue has at least 1 eigenvector.
- The collection of all eigenvectors associated to an eigenvalue λ is a subspace, and is frequently called an *eigenspace* E_λ.
- Eigenspaces are invariant subspaces of $\boldsymbol{A}$, i.e., $\boldsymbol{A}E_\lambda \subseteq E_\lambda$.
- The number of times an eigenvalue is repeated a_λ is its *algebraic multiplicity*
- The *geometric multiplicity* g_λ of an eigenvalue λ is $\dim E_\lambda$.
- *Simple* eigenvalues λ have $g_\lambda = a_\lambda = 1$.
- In general we always have $g_\lambda \leqslant a_\lambda$, so that, $1 \leqslant \sum_{\lambda \in \lambda(\boldsymbol{A})} g_\lambda \leqslant \sum_{\lambda \in \lambda(\boldsymbol{A})} a_\lambda = n$.

There are many properties of eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{C}^{n\times n}$:

- All square matrices have exactly n eigenvalues, with some possibly repeated.
- All square matrices have at least 1 eigenvector.
- Eigenvectors from *distinct* eigenvalues are linearly independent.
- Every *distinct* eigenvalue has at least 1 eigenvector.
- The collection of all eigenvectors associated to an eigenvalue λ is a subspace, and is frequently called an *eigenspace* E_λ.
- Eigenspaces are invariant subspaces of $\boldsymbol{A}$, i.e., $\boldsymbol{A}E_\lambda \subseteq E_\lambda$.
- The number of times an eigenvalue is repeated a_λ is its *algebraic multiplicity*
- The *geometric multiplicity* g_λ of an eigenvalue λ is $\dim E_\lambda$.
- *Simple* eigenvalues λ have $g_\lambda = a_\lambda = 1$.
- In general we always have $g_\lambda \leqslant a_\lambda$, so that, $1 \leqslant \sum_{\lambda\in\lambda(\boldsymbol{A})} g_\lambda \leqslant \sum_{\lambda\in\lambda(\boldsymbol{A})} a_\lambda = n$.
- Eigenvalues λ with $g_\lambda < a_\lambda$ are *defective*
- Any $\boldsymbol{A}$ with a defective eigenvalue is *defective*.

$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $\lambda(\underline{A}) = \{1\}$

$a_1 = 3, \quad g_1 = 2$

A digression: similarity transforms

Two square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are *similar* if $\exists$ an invertible $\boldsymbol{S}$ such that

$$\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}.$$

(The map $\boldsymbol{A} \mapsto \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ is a similarity transform.)

A digression: similarity transforms

Two square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are *similar* if $\exists$ an invertible $\boldsymbol{S}$ such that

$$\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}.$$

(The map $\boldsymbol{A} \mapsto \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ is a similarity transform.)

Suppose $\boldsymbol{A}$ is not defective: then with linearly independent $\boldsymbol{v}_i$, $i \in [n]$ we have the relations,

$$\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i.$$

This is equivalent to,

$$\boldsymbol{A}\boldsymbol{V} = \cancel{\boldsymbol{\Lambda}\boldsymbol{V}}, \qquad \boldsymbol{V} = \begin{pmatrix} | & & | \\ \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \\ | & & | \end{pmatrix}, \qquad \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

$\underline{V}\,\underline{\Lambda}$

A digression: similarity transforms

Two square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are *similar* if $\exists$ an invertible $\boldsymbol{S}$ such that

$$\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}.$$

(The map $\boldsymbol{A} \mapsto \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ is a similarity transform.)

Suppose $\boldsymbol{A}$ is not defective: then with linearly independent $\boldsymbol{v}_i$, $i \in [n]$ we have the relations,

$$\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i.$$

This is equivalent to,

$$\boldsymbol{A}\boldsymbol{V} = \cancel{\boldsymbol{\Lambda}\boldsymbol{V}}, \qquad \boldsymbol{V} = \begin{pmatrix} | & & | \\ \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \\ | & & | \end{pmatrix}, \qquad \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

$\underline{V}\,\underline{\Lambda}$

And since $\boldsymbol{V}$ is full-rank, then

$$\boldsymbol{A} = \cancel{\boldsymbol{V}^{-1}\boldsymbol{\Lambda}\boldsymbol{V}}, \quad \underline{V}\,\underline{\Lambda}\,\underline{V}^{-1}$$

i.e., $\boldsymbol{A}$ is similar to a diagonal matrix containing its spectrum.

Matrix diagonalizability

This motivates a key definition and consequence.

Definition

A square matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is <u>diagonalizable</u> if it is similar to a diagonal matrix.

Theorem

A square matrix $\boldsymbol{A}$ is diagonalizable iff it is not defective.

When $\boldsymbol{A}$ is not defective, it is diagonalizable via a matrix whose columns are comprised of its linearly independent eigenvectors.

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

Similarity invariances

One simple observation is that the set of eigenvalues is invariant under a(ny) similarity transform, since if $\boldsymbol{A}$ is diagonalizable with diagonal matrix $\boldsymbol{\Lambda}$, then,

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = (\boldsymbol{V}\boldsymbol{S})^{-1}\boldsymbol{\Lambda}(\boldsymbol{V}\boldsymbol{S})$$

is a diagonalization of $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ with the same diagonal matrix $\boldsymbol{\Lambda}$.

Similarity invariances

One simple observation is that the set of eigenvalues is invariant under a(ny) similarity transform, since if $\boldsymbol{A}$ is diagonalizable with diagonal matrix $\boldsymbol{\Lambda}$, then,

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = (\boldsymbol{V}\boldsymbol{S})^{-1}\,\boldsymbol{\Lambda}\,(\boldsymbol{V}\boldsymbol{S})$$

is a diagonalization of $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ with the same diagonal matrix $\boldsymbol{\Lambda}$.

More consequences follow, e.g., some familiar determinant and trace properties,

$$\det \boldsymbol{A} = \prod_{j=1}^{n} \lambda_j, \qquad \mathrm{tr}\boldsymbol{A} = \sum_{j=1}^{n} \lambda_j.$$

The above is actually true for any square matrix $\boldsymbol{A}$, defective or not.

Jordan normal form

Defective matrices certainly exist. The most common example is,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

A rather nice fact is that the above example is "spiritually" the prototypical and essentially only type of defective matrix.

Defective matrices certainly exist. The most common example is,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

A rather nice fact is that the above example is "spiritually" the prototypical and essentially only type of defective matrix.

Theorem (Jordan normal form)

Every square matrix $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ is bidiagonalizable, i.e., is similar to a bidiagonal matrix.

More specifically, let $\lambda_1, \ldots, \lambda_U$ be the unique eigenvalues of $\boldsymbol{A}$, and suppose they have algebraic and geometric multiplicities a_j and g_j, $j \in [U]$, respectively. Tnen:

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{J}\boldsymbol{V}^{-1}, \qquad \boldsymbol{V} \in \mathbb{C}^{n\times n},$$

where $\boldsymbol{V}$ is invertible, and $\boldsymbol{J}$ is bidiagonal with $\Lambda(\boldsymbol{A})$ on the diagonal, given by,

$$\boldsymbol{J} = \bigoplus_{j\in[U]} \left(\lambda_j \boldsymbol{I}_{g_j-1} \oplus \boldsymbol{J}_j\right), \qquad \boldsymbol{J}_j = \lambda_j \boldsymbol{I}_{a_j-g_j+1} + \boldsymbol{N}_{a_j-g_j+1},$$

where $\boldsymbol{N}_k$ is a $k \times k$ matrix, nonzero only on its main superdiagonal that has entries all 1.

Given $\underline{\underline{A}}$, spectrum $\lambda_1, \lambda_2, \lambda_3$

$$\begin{array}{ccc} \lambda_1 & \lambda_2 & \lambda_3 \\ \uparrow & \uparrow & \uparrow \\ a=1 & a=2 & a=2 \\ g=1 & g=2 & g=1 \end{array}$$

$\exists\ \underline{\underline{V}}$ invertible such that

$$\underline{\underline{V}}^{-1}\underline{\underline{A}}\,\underline{\underline{V}} = \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_2 & & \\ & & & \lambda_3 & 1 \\ & & & & \lambda_3 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 \underline{\underline{I}}_1 & & \\ & \lambda_2 \underline{\underline{I}}_2 & \\ & & \lambda_3 \underline{\underline{I}}_2 + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{pmatrix}$$

$$= \lambda_1 \underline{\underline{I}}_1 \oplus \lambda_2 \underline{\underline{I}}_2 \oplus \left[\lambda_3 \underline{\underline{I}}_2 + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right]$$

To diagonalizability and beyond

"Most" square matrices are diagonalizable.

This is (incredibly) powerful: a symmetric linear change of basis of the input and output spaces results in a diagonal linear operator.
The particular change of basis can be quite anisotropic (and non-isometric) in nature.

All square matrices are not diagonalizable. However, ...

$$\underline{\underline{V}} = \begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{pmatrix}$$

To diagonalizability and beyond

"Most" square matrices are diagonalizable.

This is (incredibly) powerful: a symmetric linear change of basis of the input and output spaces results in a diagonal linear operator.
The particular change of basis can be quite anisotropic (and non-isometric) in nature.

All square matrices are not diagonalizable. However, ...

- ...all square matrices are bidiagonalizable (Jordan normal form)

To diagonalizability and beyond

"Most" square matrices are diagonalizable.

This is (incredibly) powerful: a symmetric linear change of basis of the input and output spaces results in a diagonal linear operator.
The particular change of basis can be quite anisotropic (and non-isometric) in nature.

All square matrices are not diagonalizable. However, ...

- ...all square matrices are bidiagonalizable (Jordan normal form)
- ...all square matrices are unitarily triangularizable (Schur decomposition)
  While triangularizability is not as clean as diagonalizability, that there are unitary transformations accomplishing this is very attractive.

To diagonalizability and beyond

"Most" square matrices are diagonalizable.

This is (incredibly) powerful: a symmetric linear change of basis of the input and output spaces results in a diagonal linear operator.
The particular change of basis can be quite anisotropic (and non-isometric) in nature.

All square matrices are not diagonalizable. However, ...

- ...all square matrices are bidiagonalizable (Jordan normal form)
- ...all square matrices are unitarily triangularizable (Schur decomposition)
  While triangularizability is not as clean as diagonalizability, that there are unitary transformations accomplishing this is very attractive.

Are there matrices that are *unitarily* diagonalizable?

A "spectral" theorem

A seemingly unrelated algebraic definition is our starting point.

Definition

A matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is *Hermitian* if $\boldsymbol{A} = \boldsymbol{A}^*$.
(Hermitian matrices are also called *self-adjoint*, or *symmetric* when $\boldsymbol{A}$ is real-valued.)

A "spectral" theorem

A seemingly unrelated algebraic definition is our starting point.

Definition

A matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is *Hermitian* if $\boldsymbol{A} = \boldsymbol{A}^*$.
(Hermitian matrices are also called *self-adjoint*, or *symmetric* when $\boldsymbol{A}$ is real-valued.)

Theorem (Spectral Theorem for Hermitian matrices)

If $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is Hermitian, then it is unitarily diagonalizable with real eigenvalues.
(Its spectrum is real-valued, and the similarity matrix accomplishing diagonalization is unitary.)

Proof: (i) Show that $\lambda(\underline{A})$ is real.

Let $(\lambda, \underline{v})$ be any eigenpair with $\|\underline{v}\|_2^2 = 1$

$$\lambda \cdot 1 = \lambda \|\underline{v}\|_2^2 = \lambda \langle \underline{v}, \underline{v} \rangle = \langle \lambda \underline{v}, \underline{v} \rangle = \langle \underline{A}\underline{v}, \underline{v} \rangle = \langle \underline{v}, \underline{A}\underline{v} \rangle = \langle v, \lambda v \rangle$$

$= \lambda^* \langle v, v \rangle = \lambda^* \cdot 1$

$\lambda = \lambda^* \Rightarrow \lambda$ real ✓.

(ii) Show that there's a unitary similarity transform that diagonalizes $\underline{\underline{A}}$.

Idea: Let $(\lambda, \underline{v})$ be any eigenpair of $\underline{\underline{A}}$.

Define $V = \operatorname{span}\{\underline{v}\}$.

Then: $V^\perp$ is an invariant subspace of $\underline{\underline{A}}$.

$(\underline{\underline{A}} V^\perp \subseteq V^\perp)$

With this: Can consider $\underline{\underline{A}}$ "restricted" to $V^\perp$, which is an $(n-1)\times(n-1)$ Hermitian matrix.

Iterate...

<u>Details</u>: Let $(\lambda_1, \underline{v}_1)$ be an eigenpair of $\underline{\underline{A}}$

Let $\underline{\underline{Q}}_1$ be unitary, $\underline{\underline{Q}}_1 = \begin{pmatrix} | & | & & | \\ \underline{v}_1 & \underline{q}_2 & \cdots & \underline{q}_n \\ | & | & & | \end{pmatrix}$

($\underline{q}_2, \ldots \underline{q}_n$ are any ON completion of $\mathbb{C}^n$).

$(\|\underline{v}_1\|_2 = 1)$

Consider $Q_1^* A Q_1$: this is Hermitian

$$Q_1^* A Q_1 = Q_1^* \begin{pmatrix} Av_1 & Aq_2 & \cdots & Aq_n \end{pmatrix}$$

$$= Q_1^* \begin{pmatrix} \lambda_1 v_1 & Aq_2 & \cdots & Aq_n \end{pmatrix}$$

$$= \begin{pmatrix} \langle \lambda_1 v_1, v_1 \rangle & \langle Aq_2, v_1 \rangle & \cdots & \langle Aq_n, v_1 \rangle \\ \langle \lambda_1 v_1, q_2 \rangle & \vdots & & \vdots \\ \vdots & & & \\ \langle \lambda_1 v_1, q_n \rangle & \langle Aq_2, v_n \rangle & \cdots & \langle Aq_n, q_n \rangle \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ 0 & & \underline{A}_{n-1} & \\ \vdots & & & \\ 0 & & & \end{pmatrix}, \quad \underline{A}_{n-1} \text{ is Hermitian}$$

Iterate: Choose $(\lambda_2, \underline{v}_2)$ as an eigenpair of $\underline{A}_{n-1}$

Define $U_2 = \begin{pmatrix} | & | & & | \\ v_2 & q_2 & \cdots & q_n \\ | & | & & | \end{pmatrix} \in \mathbb{C}^{(n-1)\times(n-1)}$

as a unitary matrix

Define $Q_2 = \begin{pmatrix} I_1 & 0 \\ 0 & U_2 \end{pmatrix} = I_1 \oplus U_2$

Consider $Q_2^* Q_1^* A Q_1 Q_2$

$$= \begin{pmatrix} I_1 & 0 \\ 0 & U_2^* \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_{n-1} \end{pmatrix} \cdot \begin{pmatrix} I_1 & 0 \\ 0 & U_2 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & 0 \\ 0 & U_2^* A_{n-1} U_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & A_{n-2} \end{pmatrix}$$

Repeat: $\underbrace{\left(\prod_{j=1}^{n-1} Q_j\right)^*} A \underbrace{\left(\prod_{j=1}^{n-1} Q_j\right)} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$

$$Q = \prod_{j=1}^{n-1} Q_j$$

A "spectral" theorem

A seemingly unrelated algebraic definition is our starting point.

Definition

A matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is *Hermitian* if $\boldsymbol{A} = \boldsymbol{A}^*$.
(Hermitian matrices are also called *self-adjoint*, or *symmetric* when $\boldsymbol{A}$ is real-valued.)

Theorem (Spectral Theorem for Hermitian matrices)

If $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is Hermitian, then it is unitarily diagonalizable with real eigenvalues.
(Its spectrum is real-valued, and the similarity matrix accomplishing diagonalization is unitary.)

Hermitian matrices are *very* common in applications, and the spectral theorem has numerous uses.

Application I: Geometric interpretation of Hermitian matrices

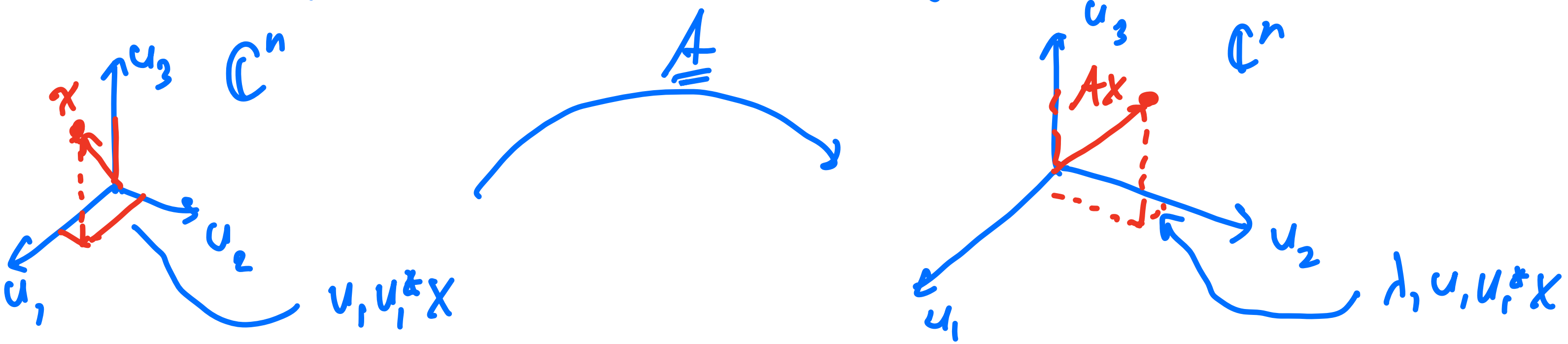
If $A \in \mathbb{C}^{n \times n}$ is unitarily diagonalizable, then it can be written as

$$A = \underbrace{U \Lambda U^*}_{} = \sum_{j=1}^{n} \lambda_j u_j u_j^*,$$

where $\{u_j\}_{j=1}^n$ are the columns of U.

$$\begin{pmatrix} \lambda_1 \underline{u}_1 & \cdots & \lambda_n \underline{u}_n \end{pmatrix} \begin{pmatrix} - & u_1^* & - \\ & \vdots & \\ - & u_n^* & - \end{pmatrix}$$

Recall: $u_j u_j^* x$ is the "vector projection" of x onto u_j



Application I: Geometric interpretation of Hermitian matrices

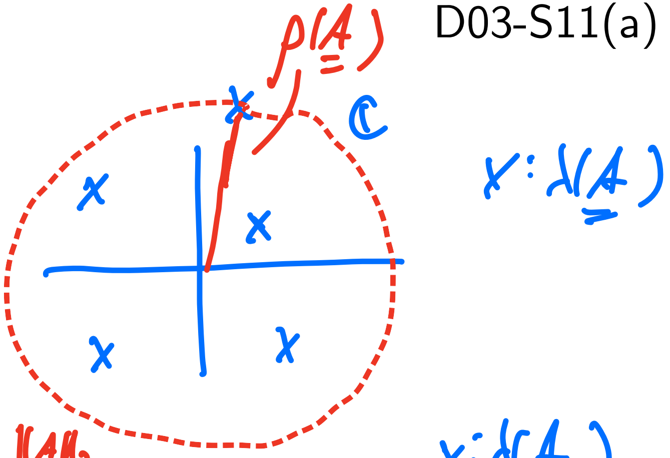
If $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is unitarily diagonalizable, then it can be written as

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^* = \sum_{j=1}^{n} \lambda_j \boldsymbol{u}_j \boldsymbol{u}_j^*,$$

where $\{\boldsymbol{u}_j\}_{j=1}^{n}$ are the columns of $\boldsymbol{U}$.

I.e., Hermitian matrices (an algebraic property) have strong geometric interpretation: they are "just" diagonal matrices in a rotated/reflected orthonormal frame.

Application II: The induced 2-norm



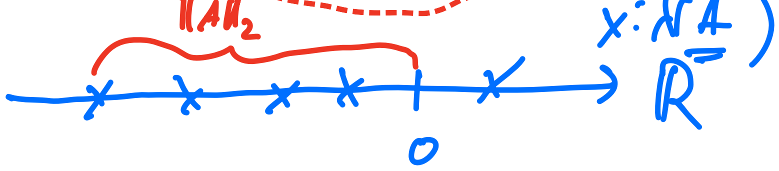
The *spectral radius* of a matrix $\boldsymbol{A}$ is

(cf. $\underline{A} = \begin{pmatrix} 0 & 10^{10} \\ 0 & 0 \end{pmatrix}$)

$$\rho(\boldsymbol{A}) := \max_{j=1,\ldots,n} |\lambda_j(\boldsymbol{A})|$$

If $\boldsymbol{A}$ is Hermitian, then $\|\boldsymbol{A}\|_2 = \rho(\boldsymbol{A})$.

This is direct from the definition of the induced 2-norm.



why? Let $\underline{x}$ satisfy $\|\underline{x}\|_2 = 1$

$$\|\underline{A}\,\underline{x}\|_2 = \Big\| \sum_{j=1}^{n} \underline{u}_j \lambda_j c_j \Big\|_2, \quad c_j = \langle x, u_j \rangle$$

$$= \Big(\sum_{j \in [n]} |c_j \lambda_j|^2 \Big)^{1/2} \leq \rho(\underline{A}) \Big(\sum_{j \in [n]} |c_j|^2 \Big)^{1/2} = \rho(\underline{A})$$

Achievable by choosing x as an eigenvector corresponding to eigenvalue achieves spectral radius

Application III: The "$\boldsymbol{A}$ norm"

A matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is Hermitian positive definite (sometimes *symmetric* positive-definite or "spd") if it's Hermitian and its (real) spectrum is strictly positive. $\lambda_1, \lambda_2 .. \lambda_n > 0$

(Respectively, positive semi-definite if the spectrum is non-negative.)

(For spd matrices, it's convention to order the spectrum $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_n$.)

Ex: $\underline{A} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is positive semi-definite, not positive definite

Application III: The "$\boldsymbol{A}$ norm"

A matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is Hermitian positive definite (sometimes *symmetric* positive-definite or "spd") if it's Hermitian and its (real) spectrum is strictly positive.

(Respectively, positive semi-definite if the spectrum is non-negative.)

Such matrices actually define a norm: $\|\boldsymbol{x}\|_{\boldsymbol{A}}^2 := \boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}$ is a norm.

$\underline{x} \mapsto \underline{x}^* \underline{\underline{A}}\, \underline{x}$ is a norm (iff $\underline{\underline{A}}$ is positive definite)

Application IV: Matrix square roots

There is a functional calculus on spd matrices.

For example, a matrix $\boldsymbol{B}$ is the square root of a matrix $\boldsymbol{A}$ if $\boldsymbol{A} = \boldsymbol{B}^2$.

Application IV: Matrix square roots

There is a functional calculus on spd matrices.

For example, a matrix $\boldsymbol{B}$ is the square root of a matrix $\boldsymbol{A}$ if $\boldsymbol{A} = \boldsymbol{B}^2$.

Example

If $\boldsymbol{A}$ is spd, compute a matrix square root of $\boldsymbol{A}$.

Idea: $\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^*$, $\underline{\underline{\Lambda}}$ has positive diagonal entries

Define $\underline{\underline{B}} = \underline{\underline{U}}\sqrt{\underline{\underline{\Lambda}}}\,\underline{\underline{U}}^*$. Then: $\underline{\underline{B}}^2 = \underline{\underline{U}}\,\underline{\underline{\Lambda}}\,\underline{\underline{U}}^* = \underline{\underline{A}}$

elementwise square root

Application IV: Matrix square roots

There is a functional calculus on spd matrices.

For example, a matrix $\boldsymbol{B}$ is the square root of a matrix $\boldsymbol{A}$ if $\boldsymbol{A} = \boldsymbol{B}^2$.

Example

If $\boldsymbol{A}$ is spd, compute a matrix square root of $\boldsymbol{A}$.

Theorem

If $\boldsymbol{A}$ is spd, then there is a unique spd square root $\boldsymbol{B}$ of $\boldsymbol{A}$, i.e., $\boldsymbol{B}^2 = \boldsymbol{A}$.

Application IV: Quadratic forms

Given a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$, the function,

$$Q_{\boldsymbol{A}}(\boldsymbol{x}) := \boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}, = \langle Ax, x \rangle$$

is a **quadratic form**, i.e., a real-valued quadratic function on $\mathbb{C}^n$. The eigendecomposition of $\boldsymbol{A}$ uniquely defines the behavior of $Q_{\boldsymbol{A}}$.

$$Q_A(x) = (U^* x)^* \Lambda (U^* x)$$

$$y = U^* x \qquad = y^* \Lambda y = \lambda_1 |y_1|^2 + \lambda_2 |y_2|^2 + \ldots$$

Application IV: Quadratic forms

Given a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$, the function,

$$Q_{\boldsymbol{A}}(\boldsymbol{x}) := \boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x},$$

is a **quadratic form**, i.e., a real-valued quadratic function on $\mathbb{C}^n$. The eigendecomposition of $\boldsymbol{A}$ uniquely defines the behavior of $Q_{\boldsymbol{A}}$.

If the following eigenvectors correspond to the positive, negative, and zero eigenvalues of $\boldsymbol{A}$, respectively,

$$\left\{\boldsymbol{v}_i^+\right\}_{i \in [n^+]}, \qquad \left\{\boldsymbol{v}_i^-\right\}_{i \in [n^-]}, \qquad \left\{\boldsymbol{v}_i^0\right\}_{i \in [n^0]},$$

where $n = n^+ + n^- + n^0$. Then clearly:

$$Q_{\boldsymbol{A}}(\boldsymbol{v}_i^+) > 0, \qquad Q_{\boldsymbol{A}}(\boldsymbol{v}_i^-) < 0, \qquad Q_{\boldsymbol{A}}(\boldsymbol{v}_i^0) = 0.$$

Application IV: Quadratic forms

Given a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$, the function,

$$Q_{\boldsymbol{A}}(\boldsymbol{x}) := \boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x},$$

is a **quadratic form**, i.e., a real-valued quadratic function on $\mathbb{C}^n$. The eigendecomposition of $\boldsymbol{A}$ uniquely defines the behavior of $Q_{\boldsymbol{A}}$.

If the following eigenvectors correspond to the positive, negative, and zero eigenvalues of $\boldsymbol{A}$, respectively,

$$\{\boldsymbol{v}_i^+\}_{i \in [n^+]}, \qquad \{\boldsymbol{v}_i^-\}_{i \in [n^-]}, \qquad \{\boldsymbol{v}_i^0\}_{i \in [n^0]},$$

where $n = n^+ + n^- + n^0$. Then clearly:

$$Q_{\boldsymbol{A}}(\boldsymbol{v}_i^+) > 0, \qquad Q_{\boldsymbol{A}}(\boldsymbol{v}_i^-) < 0, \qquad Q_{\boldsymbol{A}}(\boldsymbol{v}_i^0) = 0.$$

Generalizing this a bit:

spa

$$\left.\begin{array}{l} V^+ := \{\boldsymbol{v}_i^+\}_{i \in [n^+]}, \\ V^- := \{\boldsymbol{v}_i^-\}_{i \in [n^-]}, \\ V^0 := \{\boldsymbol{v}_i^0\}_{i \in [n^0]} \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} Q_{\boldsymbol{A}}(\boldsymbol{x}) > 0 \text{ if } \boldsymbol{x} \in V^+ \backslash \{\boldsymbol{0}\} \\ Q_{\boldsymbol{A}}(\boldsymbol{x}) < 0 \text{ if } \boldsymbol{x} \in V^- \backslash \{\boldsymbol{0}\} \\ Q_{\boldsymbol{A}}(\boldsymbol{x}) = 0 \text{ if } \boldsymbol{x} \in V^0 \end{array}\right.$$

where $\mathbb{C}^n = V^+ \oplus V^- \oplus V^0$.

Rayleigh quotients, I

A final application of Hermitian matrices is a *variational* characterization of eigenvalues. We need some buildup for this.

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be a(ny) square matrix, and let $\boldsymbol{x} \in \mathbb{C}^n \backslash \{\boldsymbol{0}\}$ be a vector.

The Rayleigh Quotient (of $\boldsymbol{A}$ at $\boldsymbol{x}$) is the (complex) scalar,

$$R_{\boldsymbol{A}}(\boldsymbol{x}) := \frac{Q_{\boldsymbol{A}}(\boldsymbol{x})}{\|\boldsymbol{x}\|_2^2} = \frac{\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^* \boldsymbol{x}} = \frac{\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}, \qquad \boldsymbol{x} \neq \boldsymbol{0}$$

Ostensibly, if $(\lambda, \boldsymbol{v})$ is an eigenpair of $\boldsymbol{A}$, then $R_{\boldsymbol{A}}(\boldsymbol{v}) = \lambda$.

$R_A(\underline{x})$ well-defined for <u>any</u> square matrix.

Rayleigh quotients, I

A final application of Hermitian matrices is a *variational* characterization of eigenvalues. We need some buildup for this.

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be a(ny) square matrix, and let $\boldsymbol{x} \in \mathbb{C}^n \backslash \{\boldsymbol{0}\}$ be a vector.

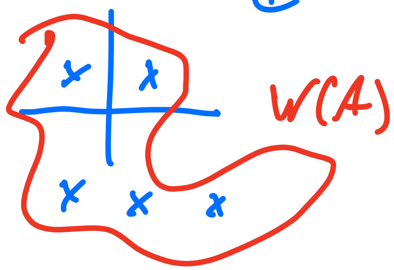
The Rayleigh Quotient (of $\boldsymbol{A}$ at $\boldsymbol{x}$) is the (complex) scalar,

$$R_{\boldsymbol{A}}(\boldsymbol{x}) := \frac{Q_{\boldsymbol{A}}(\boldsymbol{x})}{\|\boldsymbol{x}\|_2^2} = \frac{\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^* \boldsymbol{x}} = \frac{\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}, \neq \frac{\|Ax\|_2^2}{\|x\|_2^2} \qquad \boldsymbol{x} \neq \boldsymbol{0}$$

$x : N(A)$

$\mathbb{C}$

Ostensibly, if $(\lambda, \boldsymbol{v})$ is an eigenpair of $\boldsymbol{A}$, then $R_{\boldsymbol{A}}(\boldsymbol{v}) = \lambda$.



The numerical range of $\boldsymbol{A}$ is the set of all possible values of $R_{\boldsymbol{A}}$:

$$W(\boldsymbol{A}) := R_{\boldsymbol{A}} \left(\mathbb{C}^n \backslash \{\boldsymbol{0}\} \right).$$

$W(A) \subset \mathbb{C}$

One can view $W(\boldsymbol{A})$ as the image of the Rayleigh quotient over all $\mathbb{C}^n$, but also just as the image of the Rayleigh quotient over the unit sphere in $\mathbb{C}^n$.

$W(\boldsymbol{A})$ is some set in $\mathbb{C}$, regardless of the dimension n of $\boldsymbol{A}$.

Clearly we know $\lambda(\boldsymbol{A}) \subset W(\boldsymbol{A})$.

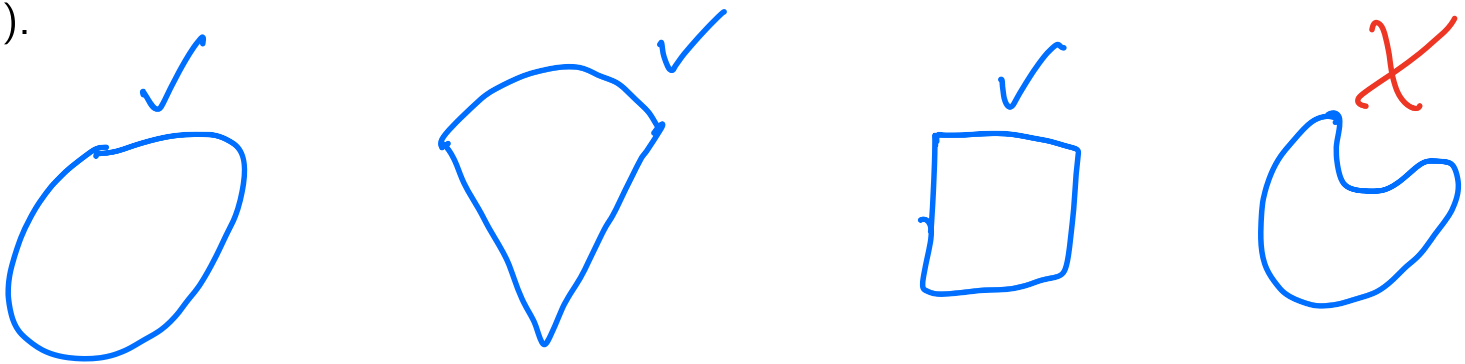
Rayleigh quotients, II

There is a rather more interesting property of the numerical range.

Theorem (Hausdorff-Toeplitz Theorem)

$W(\boldsymbol{A})$ is a compact and convex set in $\mathbb{C}$.

Compactness: $W(\boldsymbol{A})$ is the image of a compact set (unit sphere in $\mathbb{C}^n$) under a continuous function ($R_{\boldsymbol{A}}(\cdot)$).



There is a rather more interesting property of the numerical range.

Theorem (Hausdorff-Toeplitz Theorem)

$W(\boldsymbol{A})$ is a compact and convex set in $\mathbb{C}$.

Compactness: $W(\boldsymbol{A})$ is the image of a compact set (unit sphere in $\mathbb{C}^n$) under a continuous function ($R_{\boldsymbol{A}}(\cdot)$).

For certain classes of matrices, the Rayleigh quotient is a little more transparent.

For example, if $\boldsymbol{A}$ is Hermitian, then $R_{\boldsymbol{A}}(\boldsymbol{x}) \in \mathbb{R}$, so $W(\boldsymbol{A}) \subset \mathbb{R}$.

$\Downarrow$

$W(A) = [a, b] \subset \mathbb{R}$, $|a|, |b| < \infty$.

Suppose $\underline{x}$ is unit-norm: $\|x\|_2 = 1$, $R_{\underline{A}}(\underline{x}) = x^* A x = x^* U^* \Lambda U x$

Define $\underline{c} = Ux \Rightarrow R_A(x) = c^* \Lambda c = \sum_{j \in [n]} \lambda_j |c_j|^2$, $\lambda_j \in \mathbb{R}$

Order λ_j: $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$

$$\Rightarrow R_A(\underline{x}) = \sum_{j \in [n]} \lambda_j |c_j|^2 \le \sum_{j \in [n]} \lambda_n |c_j|^2 = \lambda_n \|c\|_2^2 = \lambda_n$$

Similarly: $R_A(x) \ge \lambda_1$

These are achievable: $A v_1 = \lambda_1 v_1 \Rightarrow R_A(v_1) = \lambda_1$

$A v_n = \lambda_n v_n \Rightarrow R_A(v_n) = \lambda_n$

Rayleigh quotients, II

There is a rather more interesting property of the numerical range.

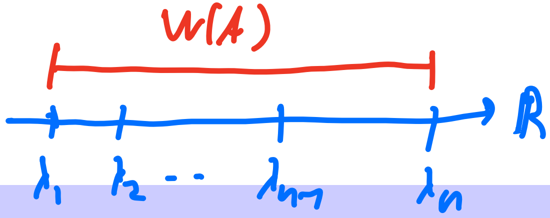
Theorem (Hausdorff-Toeplitz Theorem)

$W(\boldsymbol{A})$ is a compact and convex set in $\mathbb{C}$.

Compactness: $W(\boldsymbol{A})$ is the image of a compact set (unit sphere in $\mathbb{C}^n$) under a continuous function $(R_{\boldsymbol{A}}(\cdot))$.

For certain classes of matrices, the Rayleigh quotient is a little more transparent.

For example, if $\boldsymbol{A}$ is Hermitian, then $R_{\boldsymbol{A}}(\boldsymbol{x}) \in \mathbb{R}$, so $W(\boldsymbol{A}) \subset \mathbb{R}$.



In fact, something more precise is true

Theorem

If $\boldsymbol{A}$ is Hermitian, then

$$\lambda_{\min}(\boldsymbol{A}) \leqslant R_{\boldsymbol{A}}(\boldsymbol{x}) \leqslant \lambda_{\max}(\boldsymbol{A}), \qquad \boldsymbol{x} \in \mathbb{C}^n \backslash \{\boldsymbol{0}\}.$$

An immediate corollary: If $\boldsymbol{A}$ is Hermitian, then $W(\boldsymbol{A}) = [\lambda_{\min}(\boldsymbol{A}), \lambda_{\max}(\boldsymbol{A})]$.

Rayleigh quotients, III

Let $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ be Hermitian. Consider a subspace $V \subset \mathbb{C}^n$.

The image of the V under the Rayleigh quotient, $R_{\boldsymbol{A}}(V)$, is some subset of $W(\boldsymbol{A}) \subset \mathbb{R}$.

$$\shortparallel$$

$$\{R_A(x) \mid x \in V \setminus \{0\}\}$$

Rayleigh quotients, III

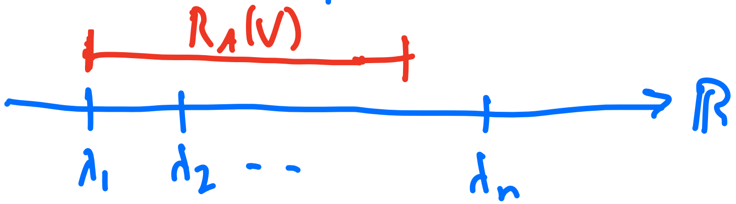
Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be Hermitian. Consider a subspace $V \subset \mathbb{C}^n$.

The image of the V under the Rayleigh quotient, $R_{\boldsymbol{A}}(V)$, is some subset of $W(\boldsymbol{A}) \subset \mathbb{R}$.

- The minimum value of $R_{\boldsymbol{A}}(V)$ is $\lambda_{\min}(\boldsymbol{A})$, occuring when V contains the minimum eigenvector. What is the <u>largest</u> possible minimum value?

If V contains $\underline{v}_1$ ($\underline{v}_1$ is eigenvector corresponding to min. eigenvalue λ_1)

$\Rightarrow R_A(\underline{v}_1) = \lambda_1 \Rightarrow \min R_A(V) = \lambda_1$



Rayleigh quotients, III

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be Hermitian. Consider a subspace $V \subset \mathbb{C}^n$.

The image of the V under the Rayleigh quotient, $R_{\boldsymbol{A}}(V)$, is some subset of $W(\boldsymbol{A}) \subset \mathbb{R}$.

- The minimum value of $R_{\boldsymbol{A}}(V)$ is $\lambda_{\min}(\boldsymbol{A})$, occuring when V contains the minimum eigenvector.
  What is the <u>largest</u> possible minimum value?
- The maximum value of $R_{\boldsymbol{A}}(V)$ is $\lambda_{\max}(\boldsymbol{A})$.
  What is the <u>smallest</u> possible maximum value?

Theorem (Courant-Fischer-Weyl "min-max")

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be Hermitian, with eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$. Then for each $1 \leqslant k \leqslant n$,

$$\lambda_k = \min_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max \overset{R_A}{\cancel{W}}_{\boldsymbol{A}}(V)$$

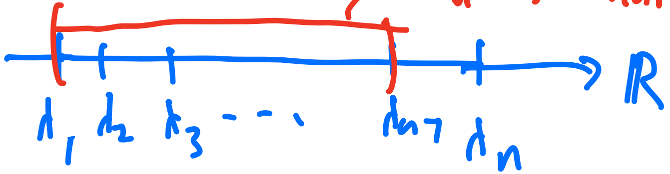
$$\lambda_k = \max_{\substack{V \subset \mathbb{C}^n \\ \dim V = n-k+1}} \min \overset{R_A}{\cancel{W}}_{\boldsymbol{A}}(V)$$

In addition, if $(\boldsymbol{u}_j)_{j=1}^n$ are the eigenvectors associated with $(\lambda_j)_{j=1}^n$, then:

- $V = \operatorname{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ *achieves the outer minimum*
- $V = \operatorname{span}\{\boldsymbol{u}_k, \ldots, \boldsymbol{u}_n\}$ *achieves the outer maximum*

$R_A(V)$ that achieves min of $\max R_A(x)$.

$k = n-1$, $\dim V = n-1$



"Proof" of 1st statement for $k=n-1$

Let V be any $(n-1)$-dim subspace

Let u_j be the eigenvector of $\underline{A}$ corresponding to λ_j.

Define $W \equiv \text{span}\{u_{n-1}, u_n\}$

$$\dim W = 2$$

$$\left.\begin{array}{l} \dim V = n-1 \\ \dim W = 2 \end{array}\right\} \dim(V \cap W) \neq 0$$

$\Rightarrow \exists\ \underline{v} \in V$ s.t. $\underline{v} = c_{n-1} u_{n-1} + c_n u_n$

c_{n-1} and c_n not both 0.

(choose $\|\underline{v}\|_2 = 1$)

$$R_A(\underline{v}) = |c_{n-1}|^2 \lambda_{n-1} + |c_n|^2 \lambda_n$$

$$\geq \lambda_{n-1}(|c_{n-1}|^2 + |c_n|^2) = \lambda_{n-1}$$

Can achieve equality if I choose V to contain u_{n-1}.

Cauchy interlacing theorem

A matrix $\boldsymbol{B}$ is a **compression** of $\boldsymbol{A}$ if $\boldsymbol{B} = \boldsymbol{Q}^* \boldsymbol{A} \boldsymbol{Q}$ for some $\boldsymbol{Q} \in \mathbb{C}^{n \times r}$ with orthonormal columns.

Cauchy interlacing theorem

A matrix $\boldsymbol{B}$ is a **compression** of $\boldsymbol{A}$ if $\boldsymbol{B} = \boldsymbol{Q}^* \boldsymbol{A} \boldsymbol{Q}$ for some $\boldsymbol{Q} \in \mathbb{C}^{n \times r}$ with orthonormal columns.

Just one consequence of the min-max theorem:

Theorem (Cauchy interlacing)

Let $\boldsymbol{B} \in \mathbb{C}^{(n-1)\times(n-1)}$ be a compression of a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$. If $\boldsymbol{A}$ has eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$, and $\boldsymbol{B}$ has eigenvalues $\mu_1, \ldots, \mu_{n-1}$, then

$$\lambda_j \leqslant \mu_j \leqslant \lambda_{j+1},$$

for all $j = 1, \ldots, n-1$.