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

Department of Mathematics, University of Utah

Fall 2026

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

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

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

Some additional terminology/properties:

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

Eigenpair properties

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

- Over $\mathbb{C}$, all $n \times n$ matrices have exactly n eigenvalues, counting algebraic multiplicity.
- All square matrices have at least 1 eigenvector.
- Eigenvectors from *distinct* eigenvalues are linearly independent.
- Every *distinct* eigenvalue has at least 1 eigenvector.
- The *eigenspace* associated with λ is the subspace $E_\lambda \coloneqq \mathsf{ker}(\mathbf{A} - \lambda \mathbf{I})$.
- Eigenspaces are invariant subspaces of $\mathbf{A}$, i.e., $\mathbf{A} E_\lambda \subseteq E_\lambda$.
- The number of times an eigenvalue is repeated a_λ is its *algebraic multiplicity*
- The *geometric multiplicity* g_λ of an eigenvalue λ is $\mathsf{dim}\, E_\lambda$.
- *Simple* eigenvalues λ have $g_\lambda = a_\lambda = 1$.
- In general we always have $g_\lambda \leq a_\lambda$, so that, $1 \leq \sum_{\lambda \in \lambda(\mathbf{A})} g_\lambda \leq \sum_{\lambda \in \lambda(\mathbf{A})} a_\lambda = n$.
- Eigenvalues λ with $g_\lambda < a_\lambda$ are *defective*
- Any $\mathbf{A}$ with a defective eigenvalue is *defective*.

A digression: similarity transforms

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

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

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

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

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

This is equivalent to,

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

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

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1},$$

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

Matrix diagonalizability

This motivates a key definition and consequence.

Definition

A square matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ is diagonalizable if it is similar to a diagonal matrix.

Theorem

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

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

Similarity invariances

One simple observation is that the set of eigenvalues is invariant under any similarity transform. If $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ is diagonalizable, then

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left(\mathbf{S}^{-1}\mathbf{V}\right)\mathbf{\Lambda}\left(\mathbf{S}^{-1}\mathbf{V}\right)^{-1}.$$

For any square matrix, defective or not, the same conclusion follows from $\mathsf{det}(\lambda\mathbf{I} - \mathbf{S}^{-1}\mathbf{A}\mathbf{S}) = \mathsf{det}(\lambda\mathbf{I} - \mathbf{A})$.

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

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

The above is actually true for any square matrix $\mathbf{A}$, 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.

Theorem (Jordan normal form)

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

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

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

where $\mathbf{V}$ is invertible and $\mathbf{J}$ is bidiagonal with $\lambda(\mathbf{A})$ on the diagonal. If $a_j = s_{j,1} + \cdots + s_{j,g_j}$, then

$$\mathbf{J} = \bigoplus_{j \in [U]} \bigoplus_{\ell=1}^{g_j} \mathbf{J}_{j,\ell}, \qquad \mathbf{J}_{j,\ell} = \lambda_j \mathbf{I}_{s_{j,\ell}} + \mathbf{N}_{s_{j,\ell}},$$

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

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.

Not all square matrices are 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 $\mathbf{A} \in \mathbb{C}^{n \times n}$ is *Hermitian* if $\mathbf{A} = \mathbf{A}^*$.
(Hermitian matrices are also called *self-adjoint*, or *symmetric* when $\mathbf{A}$ is real-valued.)

Theorem (Spectral Theorem for Hermitian matrices)

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

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

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

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 $\mathbf{A}$ is

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

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

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

Application III: The "A norm"

A matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ 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: $\|\mathbf{x}\|_{\mathbf{A}} \coloneqq \sqrt{\mathbf{x}^* \mathbf{A}\mathbf{x}}$. For positive-semidefinite $\mathbf{A}$, this expression is generally only a seminorm.

Application IV: Matrix square roots

There is a functional calculus on spd matrices.

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

Example

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

Theorem

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

Application V: Quadratic forms

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

$$Q_{\mathbf{A}}(\mathbf{x}) \coloneqq \mathbf{x}^* \mathbf{A} \mathbf{x},$$

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

Let the following vectors form an orthonormal eigenbasis, partitioned according to the positive, negative, and zero eigenvalues of $\mathbf{A}$, respectively:

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

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

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

Generalizing this a bit:

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

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 $\mathbf{A} \in \mathbb{C}^{n\times n}$ be any square matrix, and let $\mathbf{x} \in \mathbb{C}^n \backslash \{\mathbf{0}\}$ be a vector.

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

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

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

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

$$W(\mathbf{A}) \coloneqq R_{\mathbf{A}}\left(\mathbb{C}^n \backslash \{\mathbf{0}\}\right).$$

One can view $W(\mathbf{A})$ 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 $\mathbb{C}^n$.

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

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

Rayleigh quotients, II

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

Theorem (Hausdorff-Toeplitz Theorem)

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

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

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

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

In fact, something more precise is true

Theorem

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

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

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

Rayleigh quotients, III

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

The image of $V \backslash \{\mathbf{0}\}$ under the Rayleigh quotient is some subset of $W(\mathbf{A}) \subset \mathbb{R}$.

- $\min_{\mathbf{x} \in V \backslash \{\mathbf{0}\}} R_{\mathbf{A}}(\mathbf{x}) \geq \lambda_{\mathsf{min}}(\mathbf{A})$, with equality when V contains an eigenvector for $\lambda_{\mathsf{min}}(\mathbf{A})$.
  What is the <u>largest</u> possible minimum value?
- $\max_{\mathbf{x} \in V \backslash \{\mathbf{0}\}} R_{\mathbf{A}}(\mathbf{x}) \leq \lambda_{\mathsf{max}}(\mathbf{A})$, with equality when V contains an eigenvector for $\lambda_{\mathsf{max}}(\mathbf{A})$.
  What is the <u>smallest</u> possible maximum value?

The "min-max" theorem

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

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

$$\lambda_k = \min_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max_{\mathbf{x} \in V \setminus \{\mathbf{0}\}} R_{\mathbf{A}}(\mathbf{x})$$

$$\lambda_k = \max_{\substack{V \subset \mathbb{C}^n \\ \dim V = n-k+1}} \min_{\mathbf{x} \in V \setminus \{\mathbf{0}\}} R_{\mathbf{A}}(\mathbf{x})$$

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

- $V = \mathsf{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ achieves the outer minimum
- $V = \mathsf{span}\{\mathbf{u}_k, \ldots, \mathbf{u}_n\}$ achieves the outer maximum

Cauchy interlacing theorem

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

Just one consequence of the min-max theorem:

Theorem (Cauchy interlacing)

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

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

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

References

Salgado, Abner J. and Steven M. Wise (2022). *Classical Numerical Analysis: A Comprehensive Course*. Cambridge: Cambridge University Press. ISBN: 978-1-108-83770-5. DOI: 10.1017/9781108942607.

Süli, Endre and David F. Mayers (2003). *An Introduction to Numerical Analysis*. Cambridge: Cambridge University Press. ISBN: 978-0-521-00794-8. DOI: 10.1017/CBO9780511801181.

Trefethen, Lloyd N. and David Bau (1997). *Numerical Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics. ISBN: 0-89871-361-7.