Eigenvalue problems

<u>Goal:</u> Given a matrix $A \in \mathbb{C}^{n\times n}$.

Find $\lambda \in \mathbb{C}$, $v \in \mathbb{C}^n$ ($v \neq 0$) such that

$$Av = \lambda v \qquad (*)$$

λ: eigenvalue, v: eigenvector

Along direction v, matrix A acts like scalar.

When A has n linearly independent eigenvectors, we can rewrite $(*)$ for all at once

$$A \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}}_{\substack{\shortparallel \\ V \in \mathbb{C}^{n\times n}}} = \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}}_{\substack{\shortparallel \\ V}} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}}_{\substack{\shortparallel \\ \Lambda \in \mathbb{C}^{n\times n}}}$$

$$A = V \Lambda V^{-1} \qquad \leftarrow \text{eigen decomposition of } A$$

Applications: Control system, economic models, power grids, biostatistics, machine learning, ...

ex. Decoupling ODEs

$$\frac{d}{dt}\begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix} = \underbrace{\begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}}_{A = V\Lambda V^{-1}} \begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix} \qquad u(t) = (u_1(t), \cdots, u_n(t))^T \in \mathbb{R}^n$$

"n coupled ODEs with n unknowns $u_1(t), \cdots, u_n(t)$"

Solution: change to eigenvector coordinates

Let
$$\begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix} = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} c_1(t) \\ \vdots \\ c_n(t) \end{bmatrix}$$

$\leftarrow$ soln in eigenvector coordinates same basis

Then $\frac{d}{dt}\begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \dot{c}_1(t) \\ \vdots \\ \dot{c}_n(t) \end{bmatrix}$ $\qquad \dot{c}_i(t) = \frac{d}{dt} c_i(t)$, $i = 1, \cdots, n$

The system is diagonalized (decoupled)

$$\frac{d}{dt}\begin{bmatrix} c_1(t) \\ \vdots \\ c_n(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} c_1(t) \\ \vdots \\ c_n(t) \end{bmatrix}$$

$$\Leftrightarrow \quad \frac{dc_i}{dt}(t) = \lambda_i c_i(t), \quad i = 1, \cdots, n$$

$$\Leftrightarrow \quad c_i(t) = e^{\lambda_i t} c_i(0), \quad i = 1, \cdots,$$

Suppose initially $u(0) = c_1(0) v_1 + \cdots + c_n(0) v_n$

then $u(t) = e^{\lambda_1 t} c_1(0) v_1 + \cdots + e^{\lambda_n t} c_n(0) v_n$

If $Re(\lambda_i) \geq 0$, then $e^{\lambda_i t} c_i(0) v_i$ persists

If $Re(\lambda_i) < 0$, then $e^{\lambda_i t} c_i(0) v_i \to 0$ as $t \to +\infty$

$\square$

- Eigenvalue solvers must be iterative

  $A v = \lambda v \Rightarrow (A - \lambda I) v = 0 \Rightarrow \det(A - \lambda I) = 0$

  — Characteristic polynomial $p_A(\lambda) = \det(A - \lambda I)$

  Eigenvalues of $A \Leftrightarrow$ roots of $p_A(\lambda)$

  — Implication: Algorithms to find eigenvalues is very different from solving linear system $Ax = b$

  * For $Ax = b$, the solution $x = A^{-1} b$ is a rational function of a_{ij} and b_i

  * For $Ax = \lambda x$, since λ is the root of $p_A(\lambda) = \det(A - \lambda I)$,

finding λ is connected to the polynomial root finding.

λ is a root of $p(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0$ $\Longleftrightarrow$ λ is an eigenvalue of

$$\begin{bmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & & & & \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{bmatrix}$$ companion matrix of $p(\lambda)$

But by Abel-Ruffini theorem, there is no solution in radicals to general polynomial equations of degree ≥ 5

$\Rightarrow$ Any reasonable algorithm to find eigenvalues must be iterative (no direct method)

finite step solver

- Review on properties of eigenvalues / eigenvectors

— Algebraic multiplicity of λ = multiplicity of root λ of $p_A(\lambda)$

Geometric multiplicity of λ = dim of eigen space spanned by eigenvectors that satisfies $Av = \lambda v$

$= \dim V_\lambda$ $\qquad V_\lambda = \text{span}\{v \in \mathbb{C}^n : Av = \lambda v\}$

Thm: alg. mult. $\lambda \geq$ geo. mult. λ

— For any nonsingular matrix P, A and $P^{-1}AP$ are called similar, and share the same char. poly., alg./geo. mults.

— If alg. mult. $\lambda >$ geo. mult. λ, for some eigenvalue λ of A then we call A defective, otherwise nondefective.

A is nondefective $\Longleftrightarrow$ $\dim V_{\lambda_1} + \cdots + \dim V_{\lambda_r} = n$ (eigenvectors corresponding to different eigenvalues are linearly independent)

$\Leftrightarrow$ A has eigen decomposition $A = V \Lambda V^{-1}$ (diagonalizable)

where $V \in \mathbb{C}^{n \times n}$ nonsingular

and $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. $\lambda_i \in \mathbb{C}$, $i = 1, \cdots, n$

ex. Defective matrices do exist in many problems

ex. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (not singular)

$p_A(\lambda) = \det(A - \lambda I) = (\lambda - 1)^2$ $\lambda = 1$, alg. mult. $= 2$

A triangular matrix $\Rightarrow$ eigenvalues $=$ diagonal entries

$A - 1 \cdot I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. $\mathrm{rank}(A) = 1$

$\mathrm{Null}(A - 1 \cdot I) = 2 - \mathrm{rank}(A) = 1$ $\lambda = 1$ geo. mult. $= 1$

A is defective

In practice, almost no matrices are exactly defective.

(Actually, the set of $n \times n$ diagonalizable matrice is dense in $\mathbb{C}^{n \times n}$, that is, if A is defective and there exists arbitrarily small perturbation $\Delta A \in \mathbb{C}^{n \times n}$ such that $A + \Delta A$ is diagonalizable, which is usually the case when we take into account the rounding error in real implementation)

However, working with almost defective matrix is still dangerous in finding eig. decomposition.

ex. $A = \begin{pmatrix} 1+\varepsilon & 1 \\ 0 & 1-\varepsilon \end{pmatrix}$. $\lambda_1 = 1 + \varepsilon$, $\lambda_2 = 1 - \varepsilon$.

$$A v_1 = \lambda_1 v_1 \Rightarrow \quad v_1 = (1, 0)^T$$

$$A v_2 = \lambda_2 v_2 \Rightarrow \quad (1+\varepsilon) v_2^{(1)} + v_2^{(2)} = (1-\varepsilon) v_2^{(1)}$$

$$\Rightarrow \quad v_2 = (1, -2\varepsilon)^T / \sqrt{1+4\varepsilon^2}$$

$$V = \begin{pmatrix} 1 & 1/\sqrt{1+4\varepsilon^2} \\ 0 & -2\varepsilon/\sqrt{1+4\varepsilon^2} \end{pmatrix} \Rightarrow V^{-1} = \begin{pmatrix} 1 & * \\ 0 & -\sqrt{1+4\varepsilon^2}/2\varepsilon \end{pmatrix}$$

$$\Rightarrow \|V^{-1}\|_2 = O\left(\frac{1}{\varepsilon}\right)$$

These ill-conditioned similarity transformation could lead to large rounding error:

floating point matrix multiplication $\longrightarrow fl(V^{-1}AV) = V^{-1}AV + E$

where $\|E\|_2 \approx \kappa_2(V) \|A\|_2 \, \varepsilon_{mach}$

We want to work with unitary similarity for numerical stability ☑

- Schur factorization

For the purpose of finding eigenvalues, we can relax the restriction on diagonality. that is, consider a factorization of the form

$$A = Q \underset{\text{upper } \Delta}{T} Q^* .$$

the diagonal entries of T are eig. val. of A.

<u>Thm</u>: For all $A \in \mathbb{C}^{n \times n}$, $\exists$ unitary $Q \in \mathbb{C}^{n \times n}$ such that

$Q^* A Q = T$, where T is upper triangular.

Pf: induction. $n=1$ obvious.

Suppose true for $n-1$.

Let $Av = \lambda v$, $\| \|_2 = 1$.

We can find a unitary $\hat{Q}$ with 1^{st} col $= v$

$$\Rightarrow \quad \hat{Q}^* A \hat{Q} = \begin{array}{cc} \begin{array}{cc} 1 & n-1 \end{array} & \\ \begin{pmatrix} \lambda & \omega^T \\ 0 & B \end{pmatrix} & \begin{array}{c} 1 \\ n-1 \end{array} \end{array}$$

By induction, $Q_1^* B Q_1 = T_1$

$$\Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}^* \hat{Q}^* A \hat{Q} \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q_1^* \end{pmatrix} \begin{pmatrix} \lambda & \omega^T \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & Q_1^* \end{pmatrix} \begin{pmatrix} \lambda & \omega^T Q_1 \\ 0 & BQ_1 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda & \omega^T Q_1 \\ 0 & Q_1^* B Q_1 \end{pmatrix} = \begin{pmatrix} \lambda & \omega^T Q_1 \\ 0 & T_1 \end{pmatrix} =: T$$

☒

- Special case: $A^*A = AA^* \longleftarrow$ normal matrix

Thm A is normal

$\Longleftrightarrow$ A is untarily diagonalizable

$$A = Q \Lambda Q^*$$

where $Q \in \mathbb{C}^{n\times n}$ is untary $Q^*Q = I$

and $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$, $\lambda_i \in \mathbb{C}$, $i = 1, \cdots, n$

Pf: Let $A = QTQ^*$ be schur factorization of A

then $A^*A = AA^* \Longleftrightarrow T^*T = TT^*$

Lemma: If T is upper triangular and $T^*T = TT^*$

then T is diagonal. (Hint: Check off-diagonal entries row-by row)

☒

Special case: A is Hermitian $\Rightarrow$ A is normal

- Eigenvalue sensitivity:

  Many eigenvalue solvers: $V_k^{-1} A V_k \longrightarrow D$

  Question: how well do diagonal elements of a matrix approximate its eig. vals ?

Thm (Gershgorin Circle theorem)

If $A = D + N$ with $N_{ii} = 0$, $i = 1, \cdots, n$

then $\{\lambda : \lambda \text{ is an eig. val of } A\} \subset \bigcup_{i=1}^{n} D_i$

$$D_i := \{ z \in \mathbb{C} : |z - D_{ii}| \le \sum_{j=1}^{n} |N_{ij}| \}$$

or $\sum_{j=1}^{n} |N_{ji}|$

Pf: Let λ be an eig. val. of A.

We can assume $\lambda \neq D_{ii}$ $\forall i = 1, \cdots, n$

We know that $(D - \lambda I) + N$ is singular

$\Rightarrow$ $I + (D - \lambda I)^{-1} N$ is singular

$\Rightarrow$ $1 \le \| (D - \lambda I)^{-1} N \|_\infty = \sum_{j=1} \frac{|N_{i_0 j}|}{|D_{i_0 i_0} - \lambda|}$ for some i

$\Rightarrow$ $\lambda \in D_{i_0}$ ▨

Remark: It can also be shown that

if some $D_i \cap \bigcup_{j \neq i} D_j = \phi$, then D_i has exactly 1 eig. val.

Remark: This theorem is also useful in estimating eig. values.

In some methods, it is possible to show that the computed eig. vals are the exact eig. vals of $A+E$, $\|E\| \ll 1$
perturbation

We are thus interested in the following perturbation result

<u>Thm</u> (Bauer – Fike)

If μ is an eig. val of $A+E$ and

$$V^{-1}AV = D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$$

then $\displaystyle\min_{1\le i\le n} |\lambda_i - \mu| \le \kappa_p(V)\, \|E\|_p \quad \forall\, p \in [1, +\infty)$

$(\kappa_p(V) = \|V\|_p \, \|V^{-1}\|_p)$

pf: It suffices to assume $\mu \notin \{\lambda_1, \cdots, \lambda_n\}$

$$V^{-1}(A+E-\mu I)V = D-\mu I + V^{-1}EV$$
$$= (D-\mu I)\,(I + (D-\mu I)^{-1}V^{-1}EV)$$

Since $A+E-\mu I$ is singular

$\Rightarrow I + (D-\mu I)^{-1}V^{-1}EV$ is singular

$\Rightarrow 1 \le \|(D-\mu I)^{-1}V^{-1}EV\|_p \le \underbrace{\|(D-\mu I)^{-1}\|_p}_{\|(D-\mu I)^{-1}\|_p = \max_{1\le i\le n} \frac{1}{|\lambda_i - \mu|}} \|V^{-1}\|_p \, \|E\|_p \, \|V\|_p$

▨

Note: If A is normal, $\kappa_2(V) = 1$ (b.c. V is unitary)

An analogous result can be obtained via the Schur factorization

<u>Thm</u> Let $Q^* A Q = \underset{\text{diag}}{D} + \underset{\text{strict upper }\Delta}{N}$ be Schur fact. of $A \in \mathbb{C}^{n\times n}$

Let μ be an eig. val of $A+E$, and p is the smallest integer such that $|N|^p = 0$, then
$(p \le k)$

$$\min_{1 \le i \le n} |\lambda_i - \mu| \le \max\{\theta, \theta^{1/p}\}$$

where $\theta = \|E\|_2 \underline{\sum_{k=0}^{p-1} \|N\|_2^k}$

Remark: The eigenvalues of a nonnormal matrix may be sensitive to perturbation !