NLA 2021-2022
Eigenproblems

Martin Sombra

17 November 2021

Eigenvalues

Let A be an $n \times n$ matrix with complex coefficients

An element $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if there is $x \in \mathbb{C}^n \setminus \{0\}$ such that

$$A\,x = \lambda\,x$$

that is, if A is a homothecy in the direction of this n-vector

We set

$$\lambda(A) = \{\text{eigenvalues of } A\} \qquad \text{the } \textit{spectrum} \text{ of } A$$

The roots of the characteristic polynomial

The *characteristic polynomial* of the $n \times n$-matrix A is

$$\chi_A = \det(A - z\,\mathbb{1}_n) \in \mathbb{C}[z]_n,$$

where $\mathbb{C}[z]_n$ denotes the space of degree n polynomials with complex coefficients

Let $\lambda \in \mathbb{C}$. TFAE:

- $\lambda \in \lambda(A)$
- $A - \lambda\,\mathbb{1}_n$ is singular
- $\det(A - \lambda\,\mathbb{1}_n) = 0$
- λ is a root of χ_A

Hence

$$\lambda(A) = \{\lambda \in \mathbb{C} \mid \chi_A(\lambda) = 0\}.$$

Eigenvalues can be complex even if A is real!

For instance, if

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

then $\lambda(A) = \{\pm i\}$

Eigenspaces and multiplicities

For $\lambda \in \lambda(A)$, the associate *eigenspace* is

$$V_\lambda(A) = \{x \in \mathbb{C}^n \mid A\,x = \lambda\,x\}$$

It is an *invariant subsapce*:

$$A\,V_\lambda(A) \subset V_\lambda(A)$$

- The *geometric multiplicity* of λ is the dimension $\dim(V_\lambda(A))$
- The *algebraic multiplicity* of λ, denoted by $e_\lambda(A)$, is the exponent of the factor $\lambda - z$ in the irreducible factorization of χ_A

These quantities are related by

$$1 \leqslant \dim(V_\lambda(A)) \leqslant e_\lambda(A)$$

Similarities

An $n \times n$ matrix B is *similar* to A if there is a nonsingular $n \times n$ matrix S such that

$$A = S\,B\,S^{-1}$$

The matrices A and B have the same eigenvalues, and the eigenvectors of A can be read from those of B:

$$\lambda(A) = \lambda(B),$$

and y is an eigenvector of B for λ if and only if $S\,y$ is an eigenvector of A for λ

Similarities (cont.)

If $B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix, then

$$e_i = \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{array}\right] \begin{array}{c} \\ \\ \\ i \\ \\ \\ \end{array}$$

is an eigenvector of B for λ_i and so the i-th column in the matrix S

$$s_i = S\, e_i$$

is an eigenvector of A for λ_i

Decoupling

Many eigenvalue computations involve breaking down the problem into smaller ones *(decoupling)*: if

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ \mathbb{0} & A_{2,2} \end{bmatrix}$$

then $\lambda(A) = \lambda(A_{1,1}) \cup \lambda(A_{2,2})$

In particular, if A is upper triangular then its eigenvalues coincide with its diagonal entries.

The Schur decomposition

An $n \times n$ matrix Q is *unitary* if

$$Q^* = Q^{-1}$$

For reasons of numerical stability, we prefer to consider similarities given by unitary matrices

$\rightsquigarrow$ the *Schur decomposition*:

$$A = Q\,T\,Q^*$$

with Q unitary and T upper triangular

The existence of this factorization can be verified by induction:

$n = 1$:

$$Q = [1] \quad \text{and} \quad T = A$$

$n > 1$:

Take an eigenvalue $\lambda \in \lambda(A)$ and a unit eigenvector $u \in \mathbb{C}^n$ of it, and choose an $n \times (n-1)$ matrix $\widetilde{U}$ such that $U = [u \, \widetilde{U}]$ is unitary. Then

$$U^* A\, U = \begin{array}{c} \scriptstyle 1 \\ \scriptstyle n-1 \end{array} \begin{bmatrix} u^* \\ \widetilde{U}^* \end{bmatrix} A \begin{array}{cc} \scriptstyle 1 & \scriptstyle n-1 \\ \multicolumn{2}{c}{\begin{bmatrix} u & \widetilde{U} \end{bmatrix}} \end{array} = \begin{array}{c|cc} & \scriptstyle 1 & \scriptstyle n-1 \\ \hline \scriptstyle 1 & u^* A\, u & u^* A\, \widetilde{U} \\ \scriptstyle n-1 & \widetilde{U}^* A\, u & \widetilde{U}^* A\, \widetilde{U} \end{array}$$

We have that $u^* A\, u = u^* \lambda\, u = \lambda$ and $\widetilde{U}^* A\, u = \widetilde{U}^* \lambda\, u = \mathbb{0}$, and so

$$U^* A\, U = \begin{bmatrix} \lambda & \tilde{a}_{1,2} \\ \mathbb{0} & \widetilde{A}_{2,2} \end{bmatrix}$$

The Schur decomposition (cont.)

By induction, there is an $(n-1) \times (n-1)$-unitary matrix P and an $(n-1) \times (n-1)$-upper triangular matrix $\widetilde{T}$ such that

$$\widetilde{A}_{2,2} = P \, \widetilde{T} \, P^*$$

Hence

$$A = U \begin{bmatrix} \lambda & \widetilde{a}_{1,2} \\ \mathbb{0} & \widetilde{A}_{2,2} \end{bmatrix} U^* = Q \, T \, Q^*$$

with

$$Q = U \begin{bmatrix} 1 & \mathbb{0} \\ \mathbb{0} & P \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} \lambda & \widetilde{a}_{1,2} \, P \\ \mathbb{0} & \widetilde{T} \end{bmatrix}$$

Some remarks

- The Schur decomposition is *not unique*, e.g. the eigenvalues can appear in the diagonal of T in any possible order
- The columns
$$Q = [q_1 \cdots q_n]$$
are called the *Schur vectors* of A. Writing $T = [t_{i,j}]_{i,j}$, for $k = 1, \ldots, n$ we have that
$$A\, q_k = \sum_{i=1}^{k} t_{i,k}\, q_i$$
and so the linear subspace $\mathrm{span}(q_1, \ldots, q_k)$ is invariant
- When A is real symmetric, it admits a Schur decomposition $A = Q\, T\, Q^*$ with Q orthogonal and T diagonal

Diagonalization

A is *diagonalizable* if there exists a nonsingular matrix S and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ such that

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

This is equivalent to the fact that the geometric and the algebraic multiplicities of the eigenvalues of A coincide, that is,

$$\dim V_\lambda(A) = e_\lambda(A) \quad \text{for all } \lambda \in \lambda(A)$$

The Jordan decomposition

It is the factorization

$$A = S\,J\,S^{-1}$$

with J a block diagonal matrix $J = [J_1 \cdots J_q]$ where each block is of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}$$

and $n_1 + \cdots + n_q = n$

The Jordan decomposition (cont.)

This decomposition gives full information about the eigenvalues and the eigenvectors of A:

- for each i, the eigenvalue of J_i is λ_i with eigenvector e_i
- the eigenvalues of A are the λ_i's
- for each λ_i the basis of the eigenspace $V_\lambda(A)$ is given by the columns of S corresponding to the first columns of the Jordan blocks with eigenvalue λ_i

Drawbacks

- The Jordan decomposition is *not* continuous
- It is difficult to compute numerically for a non diagonalizable matrix
- It cannot be computed stably: after computing S and J, we cannot guarantee that

$$A + \delta A = S\,J\,S^{-1}$$

with δA small, because S might have a large condition number

Example 1

The Jordan form of $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is

$$J = A$$

For $\varepsilon_1 \neq \varepsilon_2$ small, the perturbed matrix

$$\widetilde{A} = \begin{bmatrix} \varepsilon_1 & 1 \\ 0 & \varepsilon_2 \end{bmatrix}$$

has two different eigenvalues ε_1 and ε_2, and so its Jordan form is

$$\widetilde{J} = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix}$$

Example 2

Let

$$A = \begin{bmatrix} 1+\varepsilon & 1 \\ 0 & 1-\varepsilon \end{bmatrix}$$

with ε small. Up to a scalar, the eigenvectors of A are

$$s_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad s_2 = \begin{bmatrix} 1 \\ -2\,\varepsilon \end{bmatrix}$$

The Jordan decomposition of A is

$$A = \begin{bmatrix} 1 & 1 \\ 0 & -2\,\varepsilon \end{bmatrix} \begin{bmatrix} 1+\varepsilon & 0 \\ 0 & 1-\varepsilon \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2\,\varepsilon} \\ 0 & \frac{-1}{2\,\varepsilon} \end{bmatrix}$$

and so $\kappa_\infty(S) \approx \varepsilon^{-1}$

Computing the eigenvectors from the Schur decomposition

Consider the Schur decomposition

$$A = Q\,T\,Q^*$$

and y an eigenvector of T for an scalar λ. Then

$$A\,Q\,y = Q\,T\,y = \lambda\,Q\,y$$

and so $Q\,y$ is an eigenvector of A with eigenvalue λ

$\rightsquigarrow$ to find the eigenvalues of A it is enough to find those of T

Computing the eigenvectors (cont.)

Suppose that $\lambda = t_{i,i}$ is *simple*, that is $e_\lambda(A) = 1$

Write the equation $(T - \lambda \mathbb{1}_n)\, y = 0$ as

$$
0 = \begin{array}{c c}
 & \begin{array}{ccc} \scriptstyle i-1 & \scriptstyle 1 & \scriptstyle n-i \end{array} \\
\begin{array}{c} \scriptstyle i-1 \\ \scriptstyle 1 \\ \scriptstyle n-i \end{array} &
\begin{bmatrix} T_{1,1} - \lambda \mathbb{1}_{i-1} & T_{1,2} & T_{1,3} \\ \mathbb{0} & 0 & T_{2,3} \\ \mathbb{0} & \mathbb{0} & T_{3,3} - \lambda \mathbb{1}_{n-i} \end{bmatrix}
\end{array}
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}
$$

$$
= \begin{bmatrix} (T_{1,1} - \lambda \mathbb{1}_{i-1})\, y_1 + T_{1,2}\, y_2 + T_{1,3}\, y_3 \\ T_{2,3}\, y_3 \\ (T_{3,3} - \lambda \mathbb{1}_{n-i})\, y_3 \end{bmatrix}
$$

Computing the eigenvectors (cont.)

Since λ is simple, both $T_{1,1} - \lambda \mathbb{1}_{i-1}$ and $T_{3,3} - \lambda \mathbb{1}_{n-i}$ are nonsingular, and so

$$y_3 = 0 \in \mathbb{R}^{i-1}$$

Set $y_2 = 1$ and compute

$$y_1 = -(T_{1,1} - \lambda \mathbb{1}_{i-1})^{-1} T_{1,2} \in \mathbb{R}^{n-i}$$

solving an upper triangular system

The resulting eigenvector is

$$y = \begin{bmatrix} -(T_{1,1} - \lambda \mathbb{1}_{i-1})^{-1} T_{1,2} \\ 1 \\ \mathbb{0} \end{bmatrix}$$

Example

For $T = \begin{bmatrix} 3 & -3 \\ 0 & 2 \end{bmatrix}$ the eigenvector for $\lambda = 2$ is computed as

$$(T - \lambda \mathbb{1}_2) \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 - 3\, y_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which gives

$$y = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Hence for a Schur decomposition $A = Q\, T\, Q^*$, the eigenvector for $\lambda = 2$ is

$$Q \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 3\, q_1 + q_2$$

Perturbation theory

Eigenvalues are continuous with respect to perturbations of the matrix, but they might have an *infinite* condition number, because in general their variation is not bounded by a linear function

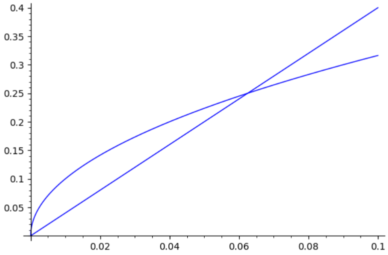
Example

Let

$$A_\varepsilon = \begin{bmatrix} 0 & 1 \\ \varepsilon & 0 \end{bmatrix}$$

with ε small

$\rightsquigarrow \lambda(A_\varepsilon) = \{\pm\varepsilon^{1/2}\}$, and $\varepsilon^{1/2}$ grows faster than $|\varepsilon|$ for $\varepsilon \to 0$



For $\varepsilon = 0$, the condition number of the eigenvalue of $\lambda = 0$ is $+\infty$

Example (cont.)

More formally, set

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A + \delta A = \begin{bmatrix} 0 & 0 \\ \varepsilon & 0 \end{bmatrix}$$

and let $\lambda = 0$ and $\lambda + \delta\lambda$ be the respective eigenvalues. Hence

$$|\delta\lambda| > c\,\frac{\|\delta A\|}{\|A\|} \approx \varepsilon$$

for any fixed $c > 0$ and ε sufficiently small

The condition number

Let A be an $n \times n$ matrix and λ a simple eigenvalue of it. The *condition number* of λ is defined as

$$\kappa(A, \lambda) = \frac{1}{y^* x}$$

where x and y are unit eigenvalues of A for λ and of A^* for $\overline{\lambda}$, respectively:

$$A x = \lambda x \quad \text{and} \quad y^* A = \lambda y^*$$

The condition number (cont.)

For a small perturbation δA we have that

$$\begin{array}{rcl} & (A+\delta A)\cdot(x+\delta x) = (\lambda+\delta\lambda)\cdot(x+\delta x) \\ - & Ax = \lambda x \\ \hline & A\cdot\delta x+\delta A\cdot x+\delta A\cdot\delta x=\lambda\cdot\delta x+\delta\lambda\cdot x+\delta\lambda\cdot\delta x \end{array}$$

If we ignore the second order terms and multiply by y^* we obtain

$$y^*\cdot A\cdot\delta x+y^*\cdot\delta A\cdot x=y^*\cdot\lambda\cdot\delta x+y^*\cdot\delta\lambda\cdot x$$

Hence up to second order

$$\delta\lambda=\frac{y^*\cdot\delta A\cdot x}{y^*\cdot x}$$

and so $|\delta\lambda| = \kappa(A,\lambda)\,\|\delta A\|$

Example (cont.)

Let

$$A_\varepsilon = \begin{bmatrix} 0 & 1 \\ \varepsilon & 0 \end{bmatrix}$$

with ε small

The eigenvalues of A and A^* for $\lambda_\varepsilon = \varepsilon^{1/2}$ are

$$x_\varepsilon = \frac{1}{(1+\varepsilon)^{1/2}} \begin{bmatrix} 1 \\ \varepsilon^{1/2} \end{bmatrix} \quad \text{and} \quad y_\varepsilon = \frac{1}{(1+\varepsilon)^{1/2}} \begin{bmatrix} \varepsilon^{1/2} \\ 1 \end{bmatrix}$$

Hence

$$\kappa(A_\varepsilon, \lambda_\varepsilon) = \frac{1}{y_\varepsilon^* \cdot x_\varepsilon} = \frac{1+\varepsilon}{2\,\varepsilon^{1/2}} \approx \frac{1}{2\,\varepsilon^{1/2}} \longrightarrow_{\varepsilon\to 0} +\infty$$

The symmetric case

When A is symmetric the condition number of a simple eigenvalue is 1: in this case $x = y$ and so

$$\kappa(A, \lambda) = \frac{1}{y^* \cdot x} = \frac{1}{\|x\|_2^2} = 1$$

Computing eigenvalues and eigenvectors

Let A be an $n \times n$ matrix and consider the *QR iteration*:

$$
\begin{aligned}
& H_0 \leftarrow Q_0^* \, A \, Q_0 \\
& \text{for } k = 1, 2, \ldots \\
& \qquad H_{k-1} = Q_k \, R_k \ \textit{(QR factorization)} \\
& \qquad H_k \leftarrow R_k \, Q_k
\end{aligned}
$$

Since $H_k = R_k \, Q_k = Q_k^* \, Q_k \, R_k \, Q_k = Q_k^* \, H_{k-1} \, Q_k$ we have that

$$A = (Q_0 \cdots Q_k) \, H_k (Q_0 \cdots Q_k)^*$$

and so H_k unitarily similar to A

In most situations

$$H_k \longrightarrow_{k \to \infty} H \qquad \text{upper triangular } \textit{(Schur form)}$$

but this is less obvious!

The power method

Fix a norm $\|\cdot\|$. The *power method* is the iteration

Choose $x_0 \in \mathbb{C}^n$ with $\|x_0\| = 1$

for $k = 0, 1, 2, \ldots$

$$y_{k+1} \leftarrow A\, x_k$$

$$x_{k+1} \leftarrow \frac{y_{k+1}}{\|y_{k+1}\|}$$

When stopping at an integer l we set

$$\widetilde{x} \leftarrow x_{l+1} \qquad \text{approximate eigenvector}$$

$$\widetilde{\lambda} \leftarrow \widetilde{x}^* A\, \widetilde{x} \qquad \text{approximate eigenvalue}$$

Error analysis

Suppose A diagonalizable, that is

$$A = S \Lambda S^{-1}$$

with

$$S = [s_1 \ \cdots \ s_n] \quad \text{and} \quad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

and suppose that

$$|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$$

Error analysis (cont.)

Write

$$x_0 = a_1\, s_1 + \cdots + a_n\, s_n$$

and suppose furthemore that $a_1 \neq 0$. Then

$$A^k x_0 = a_1\, \lambda_1^k\, s_1 + \cdots a_n\, \lambda_n^k s_n = a_1\, \lambda_1^k \Big(s_1 + \sum_{j=2}^{n} \frac{a_j}{a_1} \Big(\frac{\lambda_j}{\lambda_1}\Big)^k s_j \Big)$$

and so for $x_k = \frac{A^k x_0}{\|A^k x_0\|}$ and $\widetilde{\lambda}_k = x_k^* A\, x_k$ we have that

$$\|x_k - s_1\|, \|\widetilde{\lambda}_k - \lambda_1\| \leqslant O\Big(\Big|\frac{\lambda_2}{\lambda_1}\Big|^k \Big)$$

The number of correct digits in base b of these approximations is

$$-\log_b \|x_k - s_1\|_2, -\log_b \|\widetilde{\lambda}_k - \lambda_1\|_2 \geqslant k\, \log_b \Big(\Big|\frac{\lambda_1}{\lambda_2}\Big| \Big) + \text{constant}$$

Pros and cons

- It is based on matrix-vector multiplications
- Its speed of convergence depends on the ratio $|\lambda_1|/|\lambda_2|$
- It assumes $a_1 \neq 0$

In PageRank:

- Matrix-vector multiplications reduce to multiplications by the link matrix of the web, which is sparse
- The ratio $|\lambda_1|/|\lambda_2|$ is uniformly bounded below:

$$\frac{|\lambda_1|}{|\lambda_2|} \geqslant 0.85^{-1} \approx 1.18$$

- x_0 is chosen as the score vector of the previous web

Pros and cons (cont.)

- It only converges to a pair eigenvalue/eigenvector only if there is a dominant one
- the convergence is only *linear*

The power method does not converge in many situations, including:

- orthogonal matrices (all eigenvalues have absolute value 1)
- real matrices with complex eigenvalues (the eigenvalues come in pairs of conjugate complex numbers)