M. Sombra
2020

VIII. Eigenvalue problems

VIII.1 Eigenvalues and invariant subspaces

[GvL, §7.1.1 & §7.1.2]

Let $A \in \mathbb{C}^{n\times n}$ (possibly singular)

$\lambda \in \mathbb{C}$ is an eigenvalue if there is $x \in \mathbb{C}^n \setminus \{0\}$ s.t.

$$Ax = \lambda x$$

A is a homothecy in the direction of x.

This happens iff $A - \lambda \mathbb{1}_n$ is singular or equivalently, iff

$$\det(A - \lambda \mathbb{1}_n) = 0$$

that is iff λ is a root of the characteristic polynomial

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

degree n polynomial with complex coefficients

The spectrum of A is

$$\lambda(A) = \{\lambda \in \mathbb{C} \text{ eigenvalue of } A\}$$
$$= \{\lambda \in \mathbb{C} \mid \chi_A(\lambda) = 0\}$$

Eigenvalues can be complex even if A is real:

Ex: $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Rightarrow \lambda(A) = \{\pm i\}$

For $\lambda \in \lambda(A)$ set

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

eigenspace of λ

We have

algebraic multiplicity of λ

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

geometric multiplicity of λ

$$\chi_A(z) = \prod_{\lambda \in \lambda(A)} (\lambda - z)^{e_\lambda(A)}$$

$V_\lambda(A)$ is an <u>invariant</u> <u>subspace</u>: $A\,V_\lambda(A) \subset V_\lambda(A)$

A matrix $B \in \mathbb{C}^{n\times n}$ is <u>similar</u> to A if $\exists$ $S \in \mathbb{C}^{n\times n}$ non singular s.t

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

A and B have the <u>same eigenvalues</u>, and the <u>eigenvectors</u> of A can be <u>read from</u> those of B:

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

y is an <u>eigenvector of B for λ</u>

$$\Updownarrow$$

Sy is an <u>eigenvector of A</u> for λ

For instance, if $B = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ is diagonal then

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} i \qquad \text{eigenvector of } B \text{ for } \lambda_i$$

$$\Rightarrow \qquad s_i = S e_i \quad \text{eigenvector of } A \text{ for } \lambda_i$$

ith column of S

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

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

then $\qquad \lambda(A) = \lambda(A_{11}) \cup \lambda(A_{22})$

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

VIII.2 The Schur decomposition

[D, § 4.2] & [GvL, § 7.1.3]

A matrix $Q \in \mathbb{C}^{n\times n}$ is unitary if $\underbrace{Q^*}_{\overline{Q}^T} = Q$

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

This leads to the <u>Schur decomposition</u>:
there are Q <u>unitary</u> and <u>T upper triangular</u> s.t.

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

The <u>existence</u> of this factorization can be seen <u>by induction on n</u>:

When $n=1$ it is obvious: $Q=(1)$ & $T=A$

When $n>1$ take $\lambda \in \lambda(A)$ & $u \in \mathbb{C}^n$ unit eigenvector for λ

Choose $\tilde{v}$ st $U=(u, \tilde{v})$ $n \times n$ unitary. Then

$$U^* A U = \begin{pmatrix} u^* \\ \tilde{v}^* \end{pmatrix} A \,(u\ \tilde{v}) = \begin{pmatrix} u^* A u & u^* A \tilde{v} \\ \tilde{v}^* A u & \tilde{v}^* A \tilde{v} \end{pmatrix}$$

We have that

$$\left.\begin{aligned} u^* A u &= u^* \lambda u = \lambda \\ \tilde{v}^* A u &= \lambda \tilde{v}^* u = 0 \end{aligned}\right\} \Rightarrow U^* A U = \begin{pmatrix} \lambda & \tilde{a}_{12} \\ 0 & \tilde{A}_{22} \end{pmatrix}$$

<u>By induction</u> $\exists$ P $(n-1)\times(n-1)$ unitary
$\tilde{T}$ $(n-1)\times(n-1)$ upper triangular

s.t. $\tilde{A}_{22} = P \tilde{T} P^*$. Hence

$$A = U \begin{pmatrix} \lambda & \tilde{a}_{12} \\ 0 & \tilde{A}_{22} \end{pmatrix} U^* = \underbrace{\left(U \cdot \begin{pmatrix} 1 & \\ & P \end{pmatrix}\right)}_{Q} \cdot \underbrace{\begin{pmatrix} \lambda & \tilde{a}_{12} P \\ 0 & \tilde{T} \end{pmatrix}}_{T} \underbrace{\left(\begin{pmatrix} 1 & \\ & P^* \end{pmatrix} \cdot U^*\right)}_{Q^*}$$

is the Schur decomposition of A.

As shown in this proof, the Schur decomposition is not unique: the eigenvalues can appear in the diagonal of T in any order.

The columns

$$Q = (q_1 \cdots q_n)$$

are called Schur vectors. For $k = 1, \dots, n$

$$A q_k = \sum_{i=1}^{k} t_{ik} \, q_i \qquad T = (t_{ij})_{ij}$$

and so the linear subspace

$$\mathrm{span}(q_1, \dots, q_k)$$

is invariant.

Some matrices admit a Schur decomposition

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

with T diagonal (e.g. A symmetric). But in general, to make T "more diagonal" we need to consider non unitary singularities.

A matrix is diagonalizable if $\exists$ S non singular and Λ diagonal s.t.

$$A = S \cdot \underbrace{\Lambda}_{\mathrm{diag}(\lambda_1, \dots, \lambda_n)} \cdot S^{-1}$$

This is equivalent to the fact that the geometric and algebraic multiplicities coincide for all $\lambda \in \lambda(A)$:

$$\dim V_\lambda(A) = e_\lambda(A)$$

Indeed, writing

$$S = (s_1, \ldots, s_n)$$

the eigenspace $V_\lambda(A)$ is generated by the s_i's s.t. $\lambda_i = \lambda$.

In general, a matrix A admits a Jordan decomposition:

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

with J block diagonal:

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_q \end{pmatrix}$$

with

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

and $n_1 + \cdots + n_q = n$

It gives the full information on the eigenvalues and eigenvectors of A: for each i, the eigenvalue of J_i is $\underline{\underline{\lambda_i}}$ with eigenvector $\underline{\underline{e_i}}$.

Hence the eigenvalues of A are the λ_i's, and for each

$$\lambda \in \lambda(A)$$

The basis of the eigenspace $V_\lambda(A)$ is given by the columns of S corresponding to the first column of each of the Jordan blocks with eigenvalue λ.

Unfortunately, the Jordan decomposition of a defective (= non diagonalizable) matrix is difficult to compute numerically: it is not continuous.

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

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

For $\varepsilon_1 \neq \varepsilon_2$ small, the perturbed matrix $\tilde{A} = \begin{pmatrix} \varepsilon_1 & 1 \\ 0 & \varepsilon_2 \end{pmatrix}$ has two different eigenvalues ε_1 and ε_2 and so its Jordan form is

$$\tilde{J} = \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix}$$

Another reason is that it cannot be computed stably: after computing S and J, we cannot guarantee that

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

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

Example:

$$A = \begin{pmatrix} 1+\varepsilon & 1 \\ & 1-\varepsilon \end{pmatrix} \qquad \text{with } \varepsilon \text{ small}$$

Up to a scalar, the eigenvalues of A are

$$s_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad s_2 = \begin{pmatrix} 1 \\ -2\varepsilon \end{pmatrix}$$

The Jordan decomposition is

$$A = \begin{pmatrix} 1 & 1 \\ 0 & -2\varepsilon \end{pmatrix} \begin{pmatrix} 1+\varepsilon & 0 \\ 0 & 1-\varepsilon \end{pmatrix} \begin{pmatrix} 1 & 1/2\varepsilon \\ 0 & -1/2\varepsilon \end{pmatrix}$$

and so $\quad \kappa_2(S) \approx 1/\varepsilon$.

VIII.3 Computing eigenvectors from the Schur decomposition

[D, §4.2.1]

Let $A = Q \cdot T \cdot Q^*$

If $Ty = \lambda y$ then $AQy = QTy = \lambda Qy$

and so Qy is an eigenvector of A with eigenvalue λ.

Hence to find ~~the~~ eigenvectors of A it is enough to find those of T.

Suppose that $\lambda = t_{ii}$ has multiplicity 1 (it is <u>simple</u>)

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

$$0 = \begin{array}{cc} \begin{pmatrix} T_{11} - \lambda \mathbb{1}_{i-1} & T_{12} & T_{13} \\ 0 & 0 & T_{23} \\ 0 & 0 & T_{33} - \lambda \mathbb{1}_{n-i} \end{pmatrix} & \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \begin{array}{l} i-1 \\ 1 \\ n-i \end{array} \\ \begin{array}{ccc} i-1 & \qquad 1 & \qquad n-i \end{array} & \end{array}$$

$$= \begin{pmatrix} (T_{11} - \lambda \mathbb{1}_{i-1}) y_1 + T_{12} y_2 + T_{13} y_3 \\ T_{23} y_3 \\ (T_{33} - \lambda \mathbb{1}_{n-i}) y_3 \end{pmatrix}$$

Since λ is simple, $T_{11} - \lambda \mathbb{1}_{i-1}$ and $T_{33} - \lambda \mathbb{1}_{n-i}$ are nonsingular

$$\Rightarrow \boxed{y_3 = 0}$$

We set $\boxed{y_2 = 1}$ and so $\boxed{y_1 = -(T_{11} - \lambda \mathbb{1}_{i-1})^{-1} T_{12}}$

solving a triangular system

The eigenvalue is

$$y = \begin{pmatrix} (\lambda \mathbb{1}_{i-1} - T_{11})^{-1} T_{12} \\ 1 \\ 0 \end{pmatrix}$$

<u>Example</u> For $T = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$ the eigenvector for $\lambda = 2$ is $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \neq 0$ st

$$(T - 2\mathbb{1}_2)\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 0$$

that is, $y = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

For $A = QTQ^*$ the eigenvector for $\lambda = 2$ is

$(q_1\, q_2)$

$$Q \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3 q_1 + q_2$$

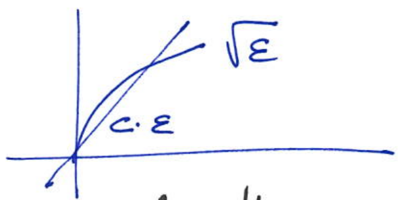
VIII.4 Perturbation theory [D, § 4.3]

Now we want to understand when eigenvalues are <u>ill-conditioned</u>, and so hard to compute.

The eigenvalues are continuous with respect to perturbations of the matrix [D, Proposition 4.4]. However, they might have an <u>infinite</u> condition number, because their variation is not bounded by a linear function.

<u>Example</u> $A_\varepsilon = \begin{pmatrix} 0 & 1 \\ \varepsilon & 0 \end{pmatrix}$ ε small

Then $\lambda(A_\varepsilon) = \{\pm\sqrt{\varepsilon}\}$
and $\pm\sqrt{\varepsilon}$ grows <u>faster</u> than $c|\varepsilon|$ $\forall c$ fixed



For $\varepsilon = 0$, the condition number of the eigenvalue $\lambda = 0$ is ∞.

More formally, set $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\delta A = \begin{pmatrix} 0 & 0 \\ \varepsilon & 0 \end{pmatrix}$,
$\lambda = 0$ $\lambda + \delta\lambda$ eigenvalue of $A + \delta A$. Then

$$\underbrace{|\delta\lambda|}_{=\sqrt{\varepsilon}} \geq c \frac{\|\delta A\|}{\|A\|} \approx \varepsilon$$

for ε small and any (fixed) $c > 0$

Let A be an $n \times n$ matrix and λ a <u>simple</u> eigenvalue.
The <u>condition number</u> of λ is

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

where x and y are <u>unit eigenvalues</u> of A and A^* respectively.

Indeed for a small perturbation δA we have that

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

Ignore the second order terms ("δ times δ") and multiply by y^*:

$$\cancel{y^* A\cdot\delta x} + y^*\delta A\, x = \cancel{y^*\lambda\cdot\delta x} + y^*\cdot\delta\lambda\cdot x$$

$$y^* A = \lambda y^*$$

and so
$$\delta\lambda = \frac{y^*\cdot\delta A\cdot x}{y^*\cdot x} \quad \text{up to second order}$$

$$\boxed{|\delta\lambda| = K(A,\lambda)\,\|\delta A\|}$$

<u>Example</u> (cont.): $\quad A_\varepsilon = \begin{pmatrix} 0 & 1 \\ \varepsilon & 0 \end{pmatrix} \quad \varepsilon > 0$ small

$$\Rightarrow \quad x_\varepsilon = \frac{1}{(1+\varepsilon)^{1/2}}\begin{pmatrix} 1 \\ \varepsilon^{1/2} \end{pmatrix} \qquad y_\varepsilon = \frac{1}{(1+\varepsilon)^{1/2}}\begin{pmatrix} \varepsilon^{1/2} \\ 1 \end{pmatrix} \quad \text{for } \lambda_\varepsilon = \varepsilon^{1/2}$$

$$K(A_\varepsilon,\lambda_\varepsilon) = \frac{1}{y_\varepsilon^* x_\varepsilon} = \frac{1+\varepsilon}{2\varepsilon^{1/2}} \approx \frac{1}{2\varepsilon^{1/2}} \xrightarrow[\varepsilon\to 0]{} \infty$$

At the other extreme, when A is symmetric the condition number of its simple eigenvalues is 1: indeed

$$y = x$$

and so

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