7 Eigenvalue problems

goal: compute some or all eigenvalues of $\mathbf{A} \in \mathbb{R}^{n\times n}$

7.1 the power method

goal: compute largest (in absolute value) eigenvalue and corresponding eigenvector

Algorithmus 7.1 (power method)

```
%input  :  $\mathbf{A} \in \mathbb{R}^{n\times n}$, $0 \neq \mathbf{x}_0 \in \mathbb{R}^n$
$\ell := 0$  ;  $\mathbf{x}_0 := \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2}$  ,  $\tilde{\lambda}_0 := \mathbf{x}_0^H \mathbf{A}\mathbf{x}_0$
repeat   {  $\mathbf{x}_{\ell+1} := \frac{\mathbf{A}\mathbf{x}_\ell}{\|\mathbf{A}\mathbf{x}_\ell\|_2}$        % approx. eigenvector
            $\tilde{\lambda}_{l+1} := \mathbf{x}_{\ell+1}^H \mathbf{A}\mathbf{x}_{\ell+1}$   % approx. eigenvalue
            $\ell := \ell + 1$
            } until sufficiently accurate                                  (7.1)
```

Satz 7.2 *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ have a basis of eigenvectors (i.e., $\mathbf{A}$ is diagonalizable) $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Let $\mathbf{x}_0 = \sum_{i=1}^n \alpha_i \mathbf{v}_1$ with $\alpha_1 \neq 0$ Then:*

(i) The $\mathbf{x}_\ell$ of Alg. 7.1 are well-defined.

(ii) $\exists C > 0$ s.t. $|\tilde{\lambda}_\ell - \lambda_1| \leq C|\frac{\lambda_2}{\lambda_1}|^\ell$, $l = 0, 1, \ldots$

Beweis: $\mathbf{x}_0 = \sum_i \alpha_i \mathbf{v}_i \;\Rightarrow\; \mathbf{A}^\ell \mathbf{x}_0 = \sum_i \alpha_i \lambda_i^\ell \mathbf{v}_i$. The assumption $\alpha_1 \neq 0 \wedge \lambda_1 \neq 0$ implies $\mathbf{A}^\ell \mathbf{x}_0 \neq 0 \forall \ell$. Inductively, this implies that $\mathbf{x}_\ell \neq 0$ for all ℓ and that $\mathbf{x}_\ell = c_\ell \mathbf{A}^\ell \mathbf{x}_0$ for $c_\ell := 1/\|\mathbf{A}^\ell \mathbf{x}_0\|_2 \neq 0$. Therefore:

$$\mathbf{x}_\ell = c_\ell \alpha_1 \lambda^\ell \left(\mathbf{v}_1 + \underbrace{\sum_{i=2}^n \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^\ell \mathbf{v}_i}_{=:\epsilon_\ell} \right). \tag{7.2}$$

The assumption $|\lambda_i| \leq |\lambda_2| < |\lambda_1| \; \forall i = 2, \ldots, n$ then implies

$$\|\epsilon_\ell\|_2 \leq \sum_{i=2}^n \left|\frac{\alpha_i}{\alpha_1}\right| \left|\frac{\lambda_i}{\lambda_1}\right|^\ell \|\mathbf{v}_i\|_2 \leq C \left|\frac{\lambda_2}{\lambda_1}\right|^\ell \text{ for suitable } C > 0. \tag{7.3}$$

For ℓ large, we have that $\|\epsilon_\ell\|_2$ is small $\Rightarrow$

$$\begin{aligned} \tilde{\lambda}_\ell &= \mathbf{x}_\ell^H \mathbf{A}\mathbf{x}_\ell \overset{\|\mathbf{x}_\ell\|_2 = 1}{=} \frac{\mathbf{x}_\ell^H \mathbf{A}\mathbf{x}_\ell}{\|\mathbf{x}_\ell\|_2^2} = \frac{(\mathbf{v}_1 + \epsilon_\ell)^H \mathbf{A}(\mathbf{v}_1 + \epsilon_\ell)}{\|\mathbf{v}_1 + \epsilon_\ell\|_2^2} = \frac{\mathbf{v}_1^H \mathbf{A}\mathbf{v}_1 + \mathbf{v}_1^H \mathbf{A}\epsilon_\ell + \epsilon_\ell^H \mathbf{A}\mathbf{v}_1 + \epsilon_\ell^H \mathbf{A}\epsilon_\ell}{\|\mathbf{v}_1 + \epsilon_\ell\|_2^2} = \\ &= \frac{\lambda_1 \|\mathbf{v}_1\|_2^2 + O(\|\epsilon_\ell\|_2)}{\|\mathbf{v}_1\|_2^2 + O(\|\epsilon_\ell\|_2)} = \lambda_1 + O(\|\epsilon_\ell\|_2) \end{aligned}$$

Hence, $|\lambda_1 - \tilde{\lambda}_\ell| \;\le\; C\|\epsilon_\ell\|_2 \;\le\; C\left|\frac{\lambda_2}{\lambda_1}\right|^\ell$. □

Bemerkung 7.3 *1. Since $\mathbf{v}_1$ is not known, the requirement $\alpha_1 \neq 0$ cannot be checked. In practice, this is not a problem since:*

- *a randomly chosen $\mathbf{x}_0$ satisfies $\alpha_1 \neq 0$ with probability 1*
- *rounding errors create a component in the direction of $\mathbf{v}_1$*

2. analogous result holds for the eigenvalue converge if λ_1 is a multiple eigenvalue

3. Algorithm 7.1 does not converge, if $\lambda_1 \neq \lambda_2$ but $|\lambda_1| = |\lambda_2|$. This case arises, e.g., when $\mathbf{A} \in \mathbb{R}^{n\times n}$ but $\mathbf{A}$ has complex eigenvalues.

4. greatest weakness of Algorithm 7.1: slow convergence if λ_1 is not well-separated from $\sigma(\mathbf{A}) \setminus \{\lambda_1\}$, i.e., $\left|\frac{\lambda_2}{\lambda_1}\right|$ is close to 1.

5. common application: estimate $\|\mathbf{A}\|_2^2 = \lambda_{max}(\mathbf{A}^H\mathbf{A})$

Folie 33

In addition to providing approximations to the largest eigenvalue, Algorithm 7.1 also yields an approximation to the corresponding eigenvector. To capture this convergence mathematically, we introduce the notion of "distance" between the spaces spanned by two vectors:

Definition 7.4 *Let $\{0\} \neq \mathcal{S} = \operatorname{span}\{\mathbf{x}\}$ and $\{0\} \neq \mathcal{T} = \operatorname{span}\{\mathbf{y}\}$. We define*

$$d(\mathcal{S},\mathcal{T}) := |\sin\varphi| = \sqrt{1-\cos^2\varphi}, \qquad \cos\varphi = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2}.$$

Bemerkung 7.5 (geometric intepretation) *φ is the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$. If $\mathbf{x} \parallel \mathbf{y}$, then $\varphi = 0$, i.e., $\mathcal{S} = \mathcal{T}$ and indeed $d(\mathcal{S},\mathcal{T}) = 0$. If $\mathbf{x} \perp \mathbf{y}$, then $d(\mathcal{S},\mathcal{T}) = 1$.* ■

The following Theorem 7.6 shows that $|\sin\angle(\mathbf{v}_1, \mathbf{x}_\ell)| \to 0$:

Satz 7.6 *Assumptions as in Theorem 7.2. Then $\exists\, C > 0$ such that*

$$d(\operatorname{span}\{\mathbf{v}_1\},\ \operatorname{span}\{\mathbf{x}_\ell\}) \;\le\; C\left|\frac{\lambda_2}{\lambda_1}\right|^\ell,\ \ell = 0,1,\ldots$$

Beweis: From (7.2), we get $\operatorname{span}\{\mathbf{x}_\ell\} = \operatorname{span}\{\mathbf{v}_1 + \epsilon_\ell\}$. Hence from (7.3) and a calculation

$$d(\operatorname{span}\{\mathbf{x}_\ell\}, \operatorname{span}\{\mathbf{v}_1\}) \le \frac{\|\epsilon_\ell\|_2}{\|\mathbf{v}_1 + \epsilon_\ell\|_2} \le C\left|\frac{\lambda_2}{\lambda_1}\right|^\ell$$

□

7.2 Inverse Iteration

goal: eigenvalue other than the largest one
observation: if $\mathbf{A}$ is invertible and $\sigma(\mathbf{A}) = \{\lambda_i \,|\, i = 1, \dots, n\}$ then $\sigma(\mathbf{A}^{-1}) = \{\frac{1}{\lambda_i} \,|\, i = 1, \dots, n\}$ i.e., the largest (in absolute value) eigenvalue of $\mathbf{A}^{-1}$ is the reciprocal of the smallest one (in absolute value) of $\mathbf{A}$.

Algorithmus 7.7 (inverse Iteration) $\ell := 0, \qquad \mathbf{x}_0 := \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2}$

```
repeat {
- solve $\mathbf{A}\tilde{\mathbf{x}}_{\ell+1} = \mathbf{x}_\ell$
- $\mathbf{x}_{\ell+1} := \frac{\tilde{\mathbf{x}}_{\ell+1}}{\|\tilde{\mathbf{x}}_{\ell+1}\|_2}$
- $\tilde{\lambda}_{\ell+1} := \mathbf{x}^H_{\ell+1}\mathbf{A}\mathbf{x}_{\ell+1}$
- $\ell := \ell + 1$
} until sufficiently accurate
```

Bemerkung 7.8 *1. If $0 < |\lambda_n| < |\lambda_{n-1}| \leq \dots \leq |\lambda_1|$, then, analogous to Theorem 7.2, one has $|\lambda_n - \tilde{\lambda}_l| \leq C \left|\frac{\lambda_n}{\lambda_{n-1}}\right|^\ell$ ⟦ exercise ⟧*

2. since a linear system is solved in each step → perform a LU-factorization of $\mathbf{A}$ at the beginning

The inverse iteration is a special case of an inverse iteration with shift:

Algorithmus 7.9 (inverse iteration with shift) *% input $\mathbf{A} \in \mathbb{R}^{n\times n}$, shift $\lambda \in \mathbb{R}$, $\mathbf{x}_0 \in \mathbb{R}^n \backslash \{0\}$*
$\ell := 0$; $\mathbf{x}_0 := \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2}$

```
repeat {
- solve $(\mathbf{A} - \lambda)\tilde{\mathbf{x}}_{\ell+1} = \mathbf{x}_\ell$
- $\mathbf{x}_{\ell+1} := \frac{\tilde{\mathbf{x}}_{\ell+1}}{\|\tilde{\mathbf{x}}_{\ell+1}\|_2}$
- $\tilde{\lambda}_{\ell+1} := \mathbf{x}^H_{\ell+1}\mathbf{A}\mathbf{x}_{\ell+1}$
- $\ell := \ell + 1$
} until sufficiently accurate
```

Satz 7.10 *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be diagonalizable; $\lambda \in \mathbb{R}$. Let the eigenvalues of $\mathbf{A}$ be numbered such that $|\lambda_1 - \lambda| \geq |\lambda_2 - \lambda| \geq \dots \geq |\lambda_{n-1} - \lambda| > |\lambda_n - \lambda| > 0$.*
Then: $\exists\, C > 0$ such that the approximation $\tilde{\lambda}_\ell$ computed by Algorithmus 7.9 satisfies:

$$|\lambda_n - \tilde{\lambda}_\ell| \leq C \left|\frac{\lambda_n - \lambda}{\lambda_{n-1} - \lambda}\right|^\ell$$

Beweis: analogous to that of Theorem 7.2. □

observation:

- inverse iteration with Shift the eigenvalue closest to the shift parameter $\lambda \to$ it is possible to seek specific eigenvalues
- the closer λ is to an eigenvalue, the faster the convergence

idea: use, in each step of the iteration, as a shift parameter λ the best available approximation to an eigenvalue → Rayleigh quotient iteration with shift $\lambda_\ell = \mathbf{x}_\ell^H \mathbf{A} \mathbf{x}_\ell$

Algorithmus 7.11 (Rayleigh quotient iteration) *% input $\mathbf{A} \in \mathbb{R}^{n\times n}$, $0 \neq \mathbf{x}_0 \in \mathbb{R}^n$, (=initial guess for eigenvector corresponding to sought eigenvalue)*

```
$\ell := 0$; $\mathbf{x}_0 := \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2}$
repeat {
- $\tilde{\lambda}_\ell := \mathbf{x}_\ell^H \mathbf{A} x_\ell$
- solve $(\mathbf{A} - \tilde{\lambda}_\ell)\tilde{\mathbf{x}}_{\ell+1} = \mathbf{x}_\ell$
- $\mathbf{x}_{\ell+1} := \frac{\tilde{\mathbf{x}}_{\ell+1}}{\|\tilde{\mathbf{x}}_{\ell+1}\|_2}$
} until sufficiently accurate
```

One expects better convergence of the Rayleigh quotient iteration than in the case of a fixed shift. One has, for example:

Satz 7.12 *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be symmetric, $\lambda \in \sigma(\mathbf{A})$ be a simple eigenvalue with corresponding eigenspace* $\operatorname{span}\{\mathbf{v}\}$*. Then: $\exists\, C > 0,\ \epsilon_0 > 0$ such that $\forall \epsilon \in (0, \epsilon_0)$: If $\mathbf{x}_0 \in \mathbb{R}^n \backslash \{0\}$ satisfies the condition $d(\operatorname{span}\{\mathbf{x}_0\},\ \operatorname{span}\{\}) < \epsilon$, then $\mathbf{x}_1$ (= one step of Algorithm 7.11) satisfies*

$$d(\operatorname{span}\{\mathbf{x}_1\},\ \operatorname{span}\{\mathbf{v}\}) \leq C\epsilon^3 \qquad \textit{and} \qquad \left| \frac{\mathbf{x}_0^H \mathbf{A} \mathbf{x}_0}{\|\mathbf{x}_0\|_2^2} - \lambda \right| \leq C\epsilon^2.$$

Beweis: See literature. Note in particular, that the result implies $\left| \frac{\mathbf{x}_1^H \mathbf{A} \mathbf{x}_1}{\|\mathbf{x}_1\|_2^2} - \lambda \right| \leq C\epsilon^6.$ □

Bemerkung 7.13 *1. analogous result holds also for general diagonalizable matrices: One then has locally quadratic (instead of cubic) convergence.*

2. iterations with variable shift are more expensive than those with fixed short for which a factorization can be amortized over several iterations.

Folie 33

7.3 error estimates–stopping criteria

7.3.1 Bauer-Fike

Question:
Relation of $\sigma(\mathbf{A})$ and $\sigma(\mathbf{A}+\Delta\mathbf{A})$?

Satz 7.14 (Bauer–Fike) *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be diagonalizable, i.e.,* $\exists\ \mathbf{T} \in \mathbb{R}^{n\times n}$ *mit* $\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \operatorname{diag}(\lambda_1,\dots,\lambda_n) =: D$*. Dann gilt: Für* $\Delta\mathbf{A} \in \mathbb{R}^{n\times n}$: $\forall\ \mu \in \sigma(\mathbf{A}+\Delta\mathbf{A}) : \min|\mu-\lambda_i| \leq \operatorname{cond}_p(\mathbf{T})\|\Delta\mathbf{A}\|_p$*, where* $\operatorname{cond}_p(\mathbf{T}) = \|\mathbf{T}\|_p\|\mathbf{T}^{-1}\|_p$ *and* $p \in [1,\infty]$ *arbitrary.*

Beweis: *Wlog* let $\mu \in \sigma(\mathbf{A}+\Delta\mathbf{A}) \setminus \sigma(\mathbf{A})$. Let $\mathbf{v}$ be an eigenvector with eigenvalue μ. Then:

$$((\mathbf{A}+\Delta\mathbf{A})-\mu I)\,\mathbf{v} = 0 \quad\Rightarrow\quad ((\mathbf{A}-\mu\mathbf{I})+\Delta\mathbf{A})\,\mathbf{v} = 0 \quad\Rightarrow\quad \left(\mathbf{I}+(\mathbf{A}-\mu)^{-1}\Delta\mathbf{A}\right)\mathbf{v} = 0 \quad\Rightarrow$$

$$\begin{aligned}
1 \quad &= \quad \frac{\|\mathbf{I}\mathbf{v}\|_p}{\|\mathbf{v}\|_p} = \frac{\|(\mathbf{A}-\mu)^{-1}\Delta\mathbf{A}\mathbf{v}\|_p}{\|\mathbf{v}\|_p} \ \leq\ \|(\mathbf{A}-\mu)^{-1}\|_p\frac{\|\Delta\mathbf{A}\mathbf{v}\|_p}{\|\mathbf{v}\|_p} \\
&\overset{\mathbf{A}=\mathbf{T}^{-1}\mathbf{D}\mathbf{T}}{\leq} \quad \|\left(\mathbf{T}^{-1}(\mathbf{D}-\mu)\mathbf{T}\right)^{-1}\|_p\|\Delta\mathbf{A}\|_p \ \leq\ \|\mathbf{T}^{-1}\|_p\|(\mathbf{D}-\mu)^{-1}\|_p\|\mathbf{T}\|_p\|\Delta\mathbf{A}\|_p \\
&= \quad \|\Delta\mathbf{A}\|_p\operatorname{cond}_p(\mathbf{T})\|\underbrace{(\mathbf{D}-\mu)^{-1}}_{diag.}\|_p = \|\Delta\mathbf{A}\|_p\operatorname{cond}_p(\mathbf{T})\max_{i=1,\dots,n}\frac{1}{|\lambda_i-\mu|} \\
&= \quad \frac{1}{\min_i(\lambda_i-\mu)}\|\Delta\mathbf{A}\|_p\operatorname{cond}_p(\mathbf{T})
\end{aligned}$$

□

Bemerkung 7.15 $\operatorname{cond}_p(\mathbf{T})$ *can be large* $\mathbf{A}$ *has eigenvectors that are close to being linearly dependent. This does* not *happen in the self-adjoint (symmetric) case:* ■

Korollar 7.16 *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be self-adjoint,* $\Delta\mathbf{A} \in \mathbb{R}^{n\times n}$*. Then:*

$$\forall\ \mu \in \sigma(\mathbf{A}+\Delta\mathbf{A}) \quad : \quad min_{\lambda\in\sigma(\mathbf{A})}|\mu-\lambda| \ \leq\ \|\Delta\mathbf{A}\|_2$$

Beweis: $\mathbf{A}$ selfadjoint $\Rightarrow \mathbf{A} = \mathbf{Q}^H\mathbf{D}\mathbf{Q}$ with $\mathbf{Q}$ orthogonal, i.e., $\operatorname{cond}_2(\mathbf{Q}) = 1$ □

7.3.2 remarks on stopping criteria

A pair $(\mathbf{x},\tilde{\lambda}) \in \mathbb{R}^n \setminus \{0\} \times \mathbb{R}$ is an eigenpair, if $\mathbf{A}\mathbf{x}-\tilde{\lambda}\mathbf{x} = 0$
hope: For $(\mathbf{x},\tilde{\lambda})$ not necessarily an eigenpair, ithe residual $\mathbf{A}\mathbf{x}-\tilde{\lambda}\mathbf{x}$ is a useful measure for the deviation from an eigenpair. We have

Satz 7.17 $\mathbf{A} \in \mathbb{R}^{n\times n}$ *diagonalizable,* $(\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \mathbf{D})$*,* $\|\mathbf{x}\|_2 = 1, \tilde{\lambda} \in \mathbb{R}$*. Set* $\mathbf{r} := \mathbf{A}\mathbf{x}-\tilde{\lambda}\mathbf{x}$*. Then:*

(i) $\min_{\lambda\in\sigma(\mathbf{A})}|\lambda-\tilde{\lambda}| \leq \operatorname{cond}_2(T)\|\mathbf{r}\|_2$

(ii) $\min_{\lambda\in\sigma(\mathbf{A})} |\lambda - \tilde\lambda| \leq \|\mathbf{r}\|_2$ *if* $\mathbf{A}$ *is selfadjoint.*

(iii) If $\tilde\lambda = \mathbf{x}^H\mathbf{A}\mathbf{x}$ *and* $\mathbf{A}$ *is selfadjoint and* $\tilde\lambda$ *sufficiently close to a simple eigenvalue of* $\mathbf{A}$*, then*

$$\min_{\lambda\in\sigma(\mathbf{A})} |\lambda - \tilde\lambda| \leq C\|\mathbf{r}\|_2^2$$

Beweis: *ad* (i): (perturbation argument)
The matrix $\mathbf{A} + \Delta\mathbf{A} := \mathbf{A} - \mathbf{r}\mathbf{x}^H$ satisfies

- $\|\Delta\mathbf{A}\|_2 = \|\mathbf{r}\|_2$
- $\tilde\lambda \in \sigma(\mathbf{A} + \Delta\mathbf{A})$, since $(\mathbf{A} + \Delta\mathbf{A})\mathbf{x} = \mathbf{A}\mathbf{x} - \mathbf{r}\underbrace{\mathbf{x}^H\mathbf{x}}_{=1} = \tilde\lambda\mathbf{x}$

The claim follows from Bauer-Fike (Theorem 7.14).
ad (ii): folgt aus (i)
ad (iii): see literature. □

$\boxed{\textbf{Folie 34}}$