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

Algorithm 7.1 (power method)

$$\begin{array}{rcl} \%\textit{input} & : & \mathbf{A} \in \mathbb{R}^{n\times n},\ 0 \neq \mathbf{x}_0 \in \mathbb{R}^n \\ \ell := 0 & ; & \mathbf{x}_0 := \dfrac{\mathbf{x}_0}{\|\mathbf{x}_0\|_2} \quad , \quad \tilde{\lambda}_0 := \mathbf{x}_0^H \mathbf{A}\mathbf{x}_0 \\ \textit{repeat} & \{ & \mathbf{x}_{\ell+1} := \dfrac{\mathbf{A}\mathbf{x}_\ell}{\|\mathbf{A}\mathbf{x}_\ell\|_2} \qquad \%\ \textit{approx. eigenvector} \\ & & \tilde{\lambda}_{l+1} := \mathbf{x}_{\ell+1}^H \mathbf{A}\mathbf{x}_{\ell+1} \quad \%\ \textit{approx. eigenvalue} \\ & & \ell := \ell + 1 \\ & & \}\ \textit{until sufficiently accurate} \end{array} \tag{7.1}$$

Theorem 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. $|\tilde{\lambda}_\ell - \lambda_1| \leq C|\frac{\lambda_2}{\lambda_1}|^\ell,\ l = 0, 1, \ldots$

Proof: $\mathbf{x}_0 = \sum_i \alpha_i \mathbf{v}_i \ \Rightarrow\ \mathbf{A}^\ell \mathbf{x}_0 = \sum_i \alpha_i \lambda_i^\ell \mathbf{v}_i$. 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, $|\lambda_1 - \tilde{\lambda}_\ell| \;\leq\; C\|\epsilon_\ell\|_2 \;\leq\; C\left|\frac{\lambda_2}{\lambda_1}\right|^\ell$. □

finis 11.DS

Remark 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 $\|\mathbf{A}\|_2^2 = \lambda_{max}(\mathbf{A}^H\mathbf{A})$

slide 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 $\{0\} \neq \mathcal{S} = \text{span}\{\mathbf{x}\}$ and $\{0\} \neq \mathcal{T} = \text{span}\{\mathbf{y}\}$. We define*

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

Remark 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 $|\sin\angle(\mathbf{v}_1, \mathbf{x}_\ell)| \to 0$:

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

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

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

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

□

7.2 Inverse Iteration

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

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

- *solve* $\mathbf{A}\tilde{\mathbf{x}}_{\ell+1} = \mathbf{x}_\ell$
- $\mathbf{x}_{\ell+1} := \frac{\tilde{\mathbf{x}}_{\ell+1}}{\|\tilde{\mathbf{x}}_{\ell+1}\|_2}$
- $\tilde{\lambda}_{\ell+1} := \mathbf{x}_{\ell+1}^H \mathbf{A}\mathbf{x}_{\ell+1}$
- $\ell := \ell + 1$

} *until sufficiently accurate*

Remark 7.8 *1. If* $0 < |\lambda_n| < |\lambda_{n-1}| \leq \cdots \leq |\lambda_1|$*, then, analogous to Theorem 7.2, one has*
$|\lambda_n - \tilde{\lambda}_l| \leq C \left|\frac{\lambda_n}{\lambda_{n-1}}\right|^\ell$ $\qquad [\![$ *exercise* $]\!]$

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

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

Algorithm 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* $(\mathbf{A} - \lambda)\tilde{\mathbf{x}}_{\ell+1} = \mathbf{x}_\ell$
- $\mathbf{x}_{\ell+1} := \frac{\tilde{\mathbf{x}}_{\ell+1}}{\|\tilde{\mathbf{x}}_{\ell+1}\|_2}$
- $\tilde{\lambda}_{\ell+1} := \mathbf{x}_{\ell+1}^H \mathbf{A}\mathbf{x}_{\ell+1}$
- $\ell := \ell + 1$

} *until sufficiently accurate*

Theorem 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* $|\lambda_1 - \lambda| \geq |\lambda_2 - \lambda| \geq \cdots \geq |\lambda_{n-1} - \lambda| > |\lambda_n - \lambda| > 0$*.*
Then: $\exists\, C > 0$ *such that the approximation* $\tilde{\lambda}_\ell$ *computed by Algorithmus 7.9 satisfies:*

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

Proof: analogous to that of Theorem 7.2. □

observation:

- inverse iteration with Shift the eigenvalue closest to the shift parameter $\lambda \rightarrow$ 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 $\rightarrow$ Rayleigh quotient iteration with shift $\lambda_\ell = \mathbf{x}_\ell^H \mathbf{A} \mathbf{x}_\ell$

Algorithm 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:

Theorem 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* $d(\operatorname{span}\{\mathbf{x}_0\},\ \operatorname{span}\{\mathbf{v}\}) < \epsilon$, *then* $\mathbf{x}_1$ *(= one step of Algorithm 7.11) satisfies*

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

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

Remark 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.*

slide 33

7.3 error estimates–stopping criteria

7.3.1 Bauer-Fike

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

Theorem 7.14 (Bauer–Fike) *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be diagonalizable, i.e., $\exists\, \mathbf{T} \in \mathbb{R}^{n\times n}$ with $\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \operatorname{diag}(\lambda_1,\dots,\lambda_n) =: \mathbf{D}$. Then: Let $\Delta\mathbf{A} \in \mathbb{R}^{n\times n}$. Then for any $\mu \in \sigma(\mathbf{A}+\Delta\mathbf{A})$ there holds $\min|\mu-\lambda_i| \le \operatorname{cond}_p(\mathbf{T})\|\Delta\mathbf{A}\|_p$, where $\operatorname{cond}_p(\mathbf{T}) = \|\mathbf{T}\|_p\|\mathbf{T}^{-1}\|_p$ and $p \in [1,\infty]$ arbitrary.*

Proof: Without loss of generality let $\mu \in \sigma(\mathbf{A}+\Delta\mathbf{A}) \setminus \sigma(\mathbf{A})$. Let $\mathbf{v}$ be an eigenvector with eigenvalue μ. Then:

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

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

□

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

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

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

Proof: $\mathbf{A}$ 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

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

(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*

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

Proof: *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): follows from (i)
ad (iii): see literature. □

slide 34

finis 12.DS

7.4 orthogonal Iteration

recall: the power iteration generates a sequence $(\mathbf{A}^\ell \operatorname{span}\{\mathbf{x}_0\})_{l=0}^\infty$ of 1-D spaces that converge to an invariant subspace of $\mathbf{A}$ (in fact, the eigenspace corresponding to the largest eigenvalue).

Idea: Perform power iteration on a k-dimensional space (described by $\mathbf{X}_0 \in \mathbb{R}^{n\times k}$)
Hope: The sequence $(\mathbf{A}^\ell \mathbf{X}_0)_{\ell=0}^\infty$ of k-dimensional spaces converges[1] to the invariant subspace that is spanned by the k dominant eigenvectors.
essential for the numerical realization:
The power iteration in Sec. 7.1 used a normalization of the vector in each space (i.e., an ONB of the space spanned by $\mathbf{A}^\ell \mathbf{x}_0$ was created). Here, an ONB of the space spanned by the columns of $\mathbf{A}^\ell \mathbf{X}_0$ is created.

Algorithm 7.18 (orthogonal iteration) *% input:* $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{X}_0 \in \mathbb{R}^{n\times k}$ *with linearly independent columns.*
$\ell := 0$
$\mathbf{X}_0 =: \mathbf{Q}_0\mathbf{R}_0$, *where* $\mathbf{Q}_0 \in \mathbb{R}^{n\times k}$ *has orthogonal columns,* $\mathbf{R} \in \mathbb{R}^{k\times k}$ *upper triangular.*
repeat {

$\mathbf{X}_{\ell+1} := \mathbf{A}\mathbf{Q}_\ell$

$\mathbf{X}_{\ell+1} =: \mathbf{Q}_{\ell+1}\mathbf{R}_{\ell+1}$ *% economy size QR-decomposition of* $\mathbf{X}_{\ell+1}$: $\mathbf{Q}_{\ell+1} \in \mathbb{R}^{n\times k}$ *has orthogonal columns* $\mathbf{R}_{\ell+1} \in \mathbb{R}^{n\times k}$ *is upper triangular*

$\ell := \ell + 1$

} **until** *sufficiently accurate*

Remark 7.19
1. *The columns of* $\mathbf{Q}_\ell$ *form an ONB of the space* $\mathbf{A}^\ell\mathcal{S}^0$ *where* $\mathcal{S}^0$ *is the space spanned by the columns of* $\mathbf{X}_0$.
2. *Orthogonalization is numerically essential: without orthogonalization one performs only* k *independent vector iterations that all converge to the same dominant eigenspace.*

Theorem 7.20 *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be diagonalizable,* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *basis of* $\mathbb{R}^n$ *of eigenvectors with corresponding eigenvalues* $\lambda_1, \ldots, \lambda_n$. *Let* $|\lambda_1| \geq |\lambda_2| \geq \ldots |\lambda_k| > |\lambda_{k+1}| \geq \cdots \geq |\lambda_n|$.
Let $\mathcal{S}^0 \subset \mathbb{R}^n$ *be the* k*-dimensional subspace spanned by the columns of* $\mathbf{X}_0 \in \mathbb{R}^{n\times k}$ *and assume* $\mathcal{S}^0 \cap \operatorname{span}\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\} = \{0\}$. *Then, there exists* $C > 0$ *such that the* k *eigenvalues* $\widetilde{\lambda}_{i,\ell}$, $i = 1, \ldots, k$, *of* $\mathbf{Q}_\ell^H \mathbf{A}\mathbf{Q}_\ell$ *satisfy*

$$\min_{\lambda\in\sigma(\mathbf{A})} |\widetilde{\lambda}_{i,\ell} - \lambda| \leq C \left|\frac{\lambda_{k+1}}{\lambda_k}\right|^\ell, \qquad i = 1, \ldots, k, \quad \ell = 0, 1, \ldots,$$

Furthermore, for any matrix $\mathbf{Q}'_\ell \in \mathbb{R}^{n\times(n-k)}$ *such that* $(\mathbf{Q}_\ell, \mathbf{Q}'_\ell)$ *is an orthogonal matrix, one has for the block matrix*

$$\mathbf{A}_\ell := (\mathbf{Q}_\ell, \mathbf{Q}'_\ell)^H \mathbf{A}(\mathbf{Q}_\ell, \mathbf{Q}'_\ell) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

[1] actually, we haven't introduced the notion of distance on the space of k-dimensional spaces, so that this statement has to remain vague

that

$$\|\mathbf{A}_{12}\|_2 \leq C \left| \frac{\lambda_{k+1}}{\lambda_k} \right|^{\ell}.$$

Proof: see literature. □

Remark 7.21 *The matrix* $(\mathbf{Q}_\ell, \mathbf{Q}'_\ell)^H \mathbf{A} (\mathbf{Q}_\ell, \mathbf{Q}'_\ell)$ *is similar to the matrix* $\mathbf{A}$*. Hence, its eigenvalues are the same as those of* $\mathbf{A}$*. Theorem 7.20 states that the eigenvalues of the block* $\mathbf{A}_{11}$ *are close to the* k *largest eigenvalues of* $\mathbf{A}$*. Theorem 7.20 also states that the block* $\mathbf{A}_{21}$ *tends to zero as* $\ell \to \infty$*. That is, the sequence of matrices* $\mathbf{A}_\ell$ *tends to block diagonal form.* ∎

slide 35

7.5 Basic QR-algorithm

A first way to understand the classical QR-algorithm (without refinements such as shift strategies) is to view it as the orthogonal iteration with starting matrix $\mathbf{X}_0 = \mathrm{I} \in \mathbb{R}^{n \times n}$:

Algorithm 7.22 (orthogonal iteration with $\mathbf{X}_0 = \mathrm{I}$) *% input:* $\mathbf{A} \in \mathbb{R}^{n\times}$, $\mathbf{X}_0 := \mathrm{I} \in \mathbb{R}^{n \times n}$
$\ell := 0$
$\mathbf{X}_0 =: \mathbf{Q}_0 \mathbf{R}_0$*, where* $\mathbf{Q}_0 \in \mathbb{R}^{n \times k}$ *has orthogonal columns,* $\mathbf{R} \in \mathbb{R}^{k \times k}$ *upper triangular.*
repeat {

$\mathbf{X}_{\ell+1} := \mathbf{A}\mathbf{Q}_\ell$

$\mathbf{X}_{\ell+1} =: \mathbf{Q}_{\ell+1}\mathbf{R}_{\ell+1}$ *% QR-decomposition of* $\mathbf{X}_{\ell+1}$ *:*
$\mathbf{Q}_{\ell+1} \in \mathbb{R}^{n \times k}$ *has orthogonal columns*
$\mathbf{R}_{\ell+1} \in \mathbb{R}^{n \times k}$ *is upper triangular*

$\ell := \ell + 1$

} **until** *sufficiently accurate*

Remark 7.23 *Algorithm 7.22 actually performs* n *orthogonal iterations simultaneously. That is, for each* $k \in \{1, \ldots, n\}$*, the first* k *columns of* $\mathbf{Q}_\ell$ *are those that would be created by the orthogonal iteration Alg. 7.18 started with* $\mathbf{X}_0 = [\mathbf{e}_1, \ldots, \mathbf{e}_k]$*. To see this, we compute with* $\mathbf{X}_0 = \mathrm{I}$

$$\mathbf{A}^\ell \mathrm{I} = \mathbf{A}^\ell \mathbf{X}_0 = \mathbf{A}^{\ell-1}\mathbf{A}\mathbf{X}_0 = \mathbf{A}^{\ell-1}\mathbf{Q}_1\mathbf{R}_1 = \mathbf{A}^{\ell-2}\mathbf{A}\mathbf{Q}_1\mathbf{R}_1 = \mathbf{A}^{\ell-2}\mathbf{Q}_2\mathbf{R}_2\mathbf{R}_1 = \cdots = \mathbf{Q}_\ell \mathbf{R}_\ell \cdots \mathbf{R}_1$$

Since the product $\mathbf{R}_\ell \cdots \mathbf{R}_1$ *is upper triangular as a product of upper triangular matrices, we see that the columns of* $\mathbf{A}^\ell[\mathbf{e}_1, \ldots, \mathbf{e}_k]$ *are linear combinations of the first* k *columns of* $\mathbf{Q}_\ell$*. Hence, for invertible* $\mathbf{A}$*, the first* k *columns of* $\mathbf{Q}_\ell$ *form an ONB of the space* $\mathbf{A}^\ell \mathcal{S}^0$*, where* $\mathcal{S}^0$ *is the space spanned by* $\mathbf{X}_0 = [\mathbf{e}_1, \ldots, \mathbf{e}_k]$*. See also Remark 7.19.* ∎

Since Alg. 7.22 performs n simultanenous orthogonal iterations (by Remark 7.23) Theorem 7.20 suggests that the sequence of matrices

$$\mathbf{A}_\ell := \mathbf{Q}_\ell^H \mathbf{A} \mathbf{Q}_\ell$$

converges to upper triangular form. Indeed, if $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ (and the technical conditions $\operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_k\} \cap \operatorname{span}\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\} = \{0\}$ for every $k \in \{1, \ldots, n\}$) then Theorem 7.20 asserts that each block $\mathbf{A}_\ell([1:k], [k+1:n])$ of $\mathbf{A}_\ell$ tend to zero. Since the matrices $\mathbf{A}_\ell$ are similar to $\mathbf{A}$, the eigenvalues of $\mathbf{A}_\ell$ and $\mathbf{A}$ coincide. Thus, the diagonal entries of the matrices $\mathbf{A}_\ell$ converge to the eigenvalues of $\mathbf{A}$.
The basic QR-algorithm creates the matrices $\mathbf{A}_\ell$ in a more efficient way than computing $\mathbf{Q}_\ell^H \mathbf{A} \mathbf{Q}$ directly. One makes the following observations:

$$\begin{aligned}
\mathbf{X}_{\ell+1} &= \mathbf{A}\mathbf{Q}_\ell = \mathbf{Q}_{\ell+1}\mathbf{R}_{\ell+1}, \\
\mathbf{A}_\ell &= \mathbf{Q}_\ell^H \mathbf{A} \mathbf{Q}_\ell = \underbrace{\mathbf{Q}_\ell^H \mathbf{Q}_{\ell+1}}_{=:\widehat{\mathbf{Q}}_{\ell+1}} \mathbf{R}_{\ell+1} \qquad \text{is "the" } QR\text{-decomposition of } \mathbf{A}_\ell, \\
\mathbf{A}_{\ell+1} &= \mathbf{Q}_{\ell+1}^H \mathbf{A} \mathbf{Q}_{\ell+1} = (\mathbf{Q}_\ell \widehat{\mathbf{Q}}_{\ell+1})^H \mathbf{A} \mathbf{Q}_{\ell+1} = \widehat{\mathbf{Q}}_{\ell+1}^H \mathbf{Q}_\ell^H \mathbf{A} \mathbf{Q}_\ell \widehat{\mathbf{Q}}_{\ell+1} = \widehat{\mathbf{Q}}_{\ell+1}^H \mathbf{A}_\ell \widehat{\mathbf{Q}}_{\ell+1} = \mathbf{R}_{\ell+1} \widehat{\mathbf{Q}}_{\ell+1}.
\end{aligned}$$

(The QR-decomposition is indeed unique for invertible matrices if one additionally fixes the sign of the diagonal entries of the R-factor—cf. Theorem 4.44.) We conclude that $\mathbf{A}_{\ell+1}$ is obtained from $\mathbf{A}_\ell$ by computing "the" QR-factorization of $\mathbf{A}_\ell$ and then multiplying the factors in reverse order. This is the classical QR-algorithm:

Algorithm 7.24 (basic form of classical QR-algorithm without shift and deflation)
% input: $\mathbf{A} \in \mathbb{R}^{n\times}$, $\mathbf{X}_0 := \mathrm{I} \in \mathbb{R}^{n\times n}$
$\ell := 0$; $\mathbf{A}_0 := \mathbf{A}$
repeat {

$\mathbf{A}_\ell =: \mathbf{Q}_\ell \mathbf{R}_\ell$ *% QR-decomposition of* $\mathbf{A}_\ell$

$\mathbf{A}_{\ell+1} := \mathbf{R}_\ell \mathbf{Q}_\ell$

$\ell := \ell + 1$

} **until** *sufficiently accurate*

Remark 7.25 *Computationally, Alg. 7.24 is still too expense as each QR-decomposition costs $O(n^3)$. In practice, $\mathbf{A}$ is brought to Hessenberg form (with cost $O(n^3)$) and then each QR-decomposition is only $O(n^2)$. See Example 4.53. This comptuationally essential: assuming that $O(n)$ QR-steps are needed to compute the n eigenvalues, the total cost are then $O(n^3) + O(n)O(n^2) = O(n^3)$. If, instead, cost $O(n^3)$ are incurred for each QR-step, then one expects the total cost to be $O(n)O(n^3) = O(n^4)$.* ∎

Remark 7.26 *In practice, the QR-algorithm is combined with the Rayleighquotient iteration idea, i.e., with suitable shifts. This improves the convergence of the algorithm.* ∎

finis 13.DS

7.6 Jacobi method(CSE)

7.6.1 Schur representation

goal: eigenvalue-revealing representation of $\mathbf{A}$
Diagonalizable matrices $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}^{-1}$ with the diagonal matrix $\mathbf{D}$. This is an eigenvalue-revealing representation. However, if $\mathrm{cond}(\mathbf{T})$ is large, then this representation is numerically not advisable. In this case, an alternative is the Schur form

Theorem 7.27 *Let $\mathbf{A} \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix $\mathbf{Q} \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $\mathbf{R}$ with $\mathbf{A} = \mathbf{Q}\mathbf{R}\mathbf{Q}^H$. The diagonal entries of $\mathbf{R}$ are the eigenvalues (according to multiplicity) of $\mathbf{A}$.*

Proof: We prove the theorem by induction on n. For $n = 1$ the theorem is obviously true. Suppose it is true for all matrices in $\mathbb{R}^{(n-1)\times(n-1)}$. Let $\mathbf{v} \in \mathbb{R}^n$ be an eigenvector of $\mathbf{A}$, i.e., $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Let the columns of $\mathbf{V}' \in \mathbb{C}^{n\times(n-1)}$ be such that $\mathbf{V} := (\mathbf{v}, \mathbf{V}')$ is unitary (i.e., the columns of $\mathbf{V}'$ are an ONB of the orthogonal complement of $\mathrm{span}\{\mathbf{v}\}$).

$$\mathbf{V}^H\mathbf{A}\mathbf{V} = \begin{pmatrix} \lambda & \mathbf{w}^\top \\ 0 & \mathbf{C} \end{pmatrix}, \qquad \mathbf{w} \in \mathbb{R}^{n-1}, \quad \mathbf{C} \in \mathbb{R}^{(n-1)\times(n-1)}.$$

By the induction hypothesis, there is a unitary $\mathbf{Q} \in \mathbb{R}^{(n-1)\times(n-1)}$ such that $\mathbf{Q}^H\mathbf{C}\mathbf{Q} = \mathbf{R}'$ is upper triangular. Then

$$\begin{pmatrix} 1 & 0 \\ 0 & \mathbf{Q} \end{pmatrix}^H \mathbf{V}^H\mathbf{A}\mathbf{V} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{Q} \end{pmatrix}$$

is upper triangular. Thus we have obtained the desired Schur decomposition for $\mathbf{A} \in \mathbb{C}^{n\times n}$. □

7.6.2 Jacobi method

The QR-method for eigenvalue computations is based on the idea of finding a sequence of orthogonal matrices $\mathbf{Q}_n$ such that the $\mathbf{Q}_n^\top\mathbf{A}\mathbf{Q}_n$ converges to upper triangular form. Since these are similarity transformations of $\mathbf{A}$, the diagonal entries of the upper triangular matrix contains the eigenvalues. If the entries in lower part are small, then these diagonal entries are indeed good approximations to the eigenvalues:

Exercise 7.28 *Consider a matrix $\mathbf{R} + \Delta\mathbf{A}$ where $\mathbf{R}$ is upper triangular. Show, using Theorem 7.14 that for each $\lambda \in \sigma(\mathbf{R} + \Delta\mathbf{A})$ there is a diagonal entry $\mathbf{R}_{ii}$*

$$|\lambda - \mathbf{R}_{ii}| \le C\|\Delta\mathbf{A}\|_2,$$

where the constant C depends on $\mathbf{R}$ but is independent of $\Delta\mathbf{A}$. ∎

A simpler form than the QR-method is Jacobi's method, which constructs the $\mathbf{Q}_n$ by Givens rotations. Recall the definition of Givens rotations $\mathbf{G}(i, j, \theta)$ of Section 4.6.4. We also introduce for a matrix $\mathbf{A}$

$$\mathrm{off}(\mathbf{A})^2 := \sum_{\substack{i,j \\ i\neq j}} |\mathbf{A}_{ij}|^2 = \|\mathbf{A}\|_F^2 - \sum_{i=1}^{n} |\mathbf{A}_{ii}|^2. \tag{7.4}$$

We consider *symmetric* matrices $\mathbf{A}$. The basic step of the Jacobi eigenvalue procedure consists of three steps:

1. select a pair (i, j) with $1 \leq i < j \leq n$
2. select θ such that (we write again $c = \cos\theta$, $s = \sin\theta$)
$$\begin{pmatrix} \mathbf{B}_{ii} & \mathbf{B}_{ij} \\ \mathbf{B}_{ji} & \mathbf{B}_{jj} \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}^\top \begin{pmatrix} \mathbf{A}_{ii} & \mathbf{A}_{ij} \\ \mathbf{A}_{ji} & \mathbf{A}_{jj} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \tag{7.5}$$
is diagonal
3. overwrite $\mathbf{A}$ with $\mathbf{B} = \mathbf{G}(i, j, \theta)^\top \mathbf{A}\mathbf{G}(i, j, \theta)$

In other words: one makes a similarity transformation of $\mathbf{A}$ with a Givens rotation in such a way that the entries (i, j) and (j, i) of $\mathbf{A}$ are annihilated. We now show that the transformed matrix has a smaller off-diagonal part:

Lemma 7.29 *Let* $\mathbf{A}$ *be symmetric. Let* $\mathbf{B} := \mathbf{G}(i, j, \theta)^\top \mathbf{A}\mathbf{G}(i, j, \theta)$*, where* θ *is chosen such that* $\mathbf{B}_{ij} = \mathbf{B}_{ji} = 0$*. Then*
$$\mathrm{off}(\mathbf{B})^2 = \mathrm{off}(\mathbf{A})^2 - 2\mathbf{A}_{ij}^2.$$

Proof: We consider the transformation (7.5). Since the Frobenius norm is invariant under orthogonal transformations, we have
$$\mathbf{A}_{ii}^2 + \mathbf{A}_{jj}^2 + 2\mathbf{A}_{ij}^2 = \mathbf{B}_{ii}^2 + \mathbf{B}_{jj}^2 + 2\mathbf{B}_{ij}^2 = \mathbf{B}_{ii}^2 + \mathbf{B}_{jj}^2. \tag{7.6}$$
Hence, we get
$$\begin{aligned}
\mathrm{off}(\mathbf{B})^2 &= \|\mathbf{B}\|_F^2 - \sum_{k=1}^{n} |\mathbf{B}_{kk}|^2 \\
&= \|\mathbf{A}\|_F^2 - \sum_{k \notin \{i,j\}} |\mathbf{B}_{kk}|^2 - |\mathbf{B}_{ii}|^2 - |\mathbf{B}_{jj}|^2 \\
&\overset{\text{only rows/colums } i,\, j \text{ are touched}}{=} \|\mathbf{A}\|_F^2 - \sum_{k \notin \{i,j\}} |\mathbf{A}_{kk}|^2 - |\mathbf{B}_{ii}|^2 - |\mathbf{B}_{jj}|^2 \\
&= \|\mathbf{A}\|_F^2 - \sum_{k=1}^{n} |\mathbf{A}_{kk}|^2 + |\mathbf{A}_{ii}|^2 + |\mathbf{A}_{jj}|^2 - |\mathbf{B}_{ii}|^2 - |\mathbf{B}_{jj}|^2 \\
&\overset{(7.6)}{=} \mathrm{off}(\mathbf{A})^2 + |\mathbf{A}_{ii}|^2 - 2|\mathbf{A}_{ij}|^2.
\end{aligned}$$

□

Lemma 7.29 suggests that one should select the pair i, j such that $|\mathbf{A}_{ij}|$ is as large as possible. That is, taking the largest possible off-diagonal entry yields, with $N = n(n-1)/2$
$$\mathrm{off}(\mathbf{A})^2 \leq N(\mathbf{A}_{ij}^2 + \mathbf{A}_{ji}^2) \tag{7.7}$$
and therefore
$$\mathrm{off}(\mathbf{B})^2 \overset{\text{Lemma 7.29}}{=} \mathrm{off}(\mathbf{A})^2 - \left(|\mathbf{A}_{ij}|^2 + |\mathbf{A}_{ji}|^2\right) \overset{(7.7)}{\leq} \left(1 - \frac{1}{N}\right) \mathrm{off}(\mathbf{A})^2.$$

Thus, the Jacobi method converges to upper triangular form.

Remark 7.30 *1. Searching the largest off-diagonal entry incurs large costs. Practically, one therefore simply loops through the off-diagonal entries of* $\mathbf{A}$. $\rightarrow$ *"cyclic Jacobi" method.*

2. The convergence is linear. However, the asymptotic convergence is actually quadratic, i.e., for the k-th matrix $\mathbf{A}^{(k)}$ *one has* $\operatorname{off}(\mathbf{A}^{(k+N)})^2 \leq C \operatorname{off}(\mathbf{A}^{(k)})^4$.

3. The Jacobi method is not competitive with the QR-algorithm in general. However, if $\mathbf{A}$ *is already close to diagonal, then it is an option.*

4. Variants exist that produce the SVD of $\mathbf{A}$.

■

7.7 QR-algorithm with shift (CSE)

goal: convergence acceleration of QR-algorithm using shifts.
mathematical background: Implicitly the QR-algorithm with shift performs an inverse iteration for $\mathbf{A}^H$ so that choosing Rayleigh quotients as shift leads to rapid convergence.
So far, we assumed $\mathbf{A}$ to be real (although this is by no means essential). Since we want to allow complex shifts, we allow $\mathbf{A}$ to be complex. We note that the concept of QR-factorizations also holds for complex matrices.[2]
A generalization of the basic QR-algorithm is the QR-algorithm with shift:

Algorithm 7.31 (QR-Algorithm with shift) $\ell := 0 \qquad \mathbf{A}_0 := 0$

repeat {
- *choose shift* $\mu^{(\ell)}$
- $\mathbf{A}_\ell - \mu^{(\ell)} := \mathbf{Q}_{\ell+1}\mathbf{R}_{\ell+1}$
- $\mathbf{A}_{\ell+1} := \mathbf{R}_{\ell+1}\mathbf{Q}_{\ell+1} + \mu^{(\ell)}$

} *until sufficiently accurate*

Exercise 7.32 *Check that* $\mathbf{A}_\ell$ *and* $\mathbf{A}_{\ell+1}$ *are similar and hence have the same eigenvalues.* ■

Let us consider the case that the shift μ^ℓ is an eigenvalue of $\mathbf{A}_\ell$. Then, $\mathbf{A}_\ell - \mu^\ell$ is singular and therefore $\mathbf{R}_{\ell+1}$ has a zero on it diagonal. Let this be the $\mathbf{R}_{nn}$ entry (this would happen, if a QR-factorization with pivoting is employed). Then, matrix $(\mathbf{R}_{\ell+1}\mathbf{Q}_{\ell+1})(n,:)$ is zero so that $\mathbf{A}_{\ell+1}(n,:) = (\mathbf{R}_{\ell+1}\mathbf{Q}_{\ell+1} + \mu^\ell)(n,:) = (0,0,\ldots,0,\mu^\ell)$. That is, $\mathbf{A}_{\ell+1}$ has "block triangular form" and μ^ℓ is identified as one eigenvalue of $\mathbf{A}_{\ell+1}$. We also note that "deflation" is then possible, i.e., one could continue with the $(n-1)\times(n-1)$ submatrix $\mathbf{A}_{\ell+1}(1:n-1,1:n-1)$ instead of $\mathbf{A}_{\ell+1}$. In general one would expect that a shift close to an eigenvalue leads to a *small* off-diagonal entry $\mathbf{A}_{\ell+1}(n,n-1)$ so that, by Theorem 7.14, the entry $\mathbf{A}_{\ell+1}(n,n)$ is a good approximation to an eigenvalue.

[2] in fact, the eigenvalue algorithms are probably better understood by viewing $\mathbf{A} \in \mathbb{C}^{n\times n}$ and specializing to real matrices if necessary

Lemma 7.33 *Let the shifts $\mu^{(\ell)}$ be such that $\mu^{(\ell)} \notin \sigma(\mathbf{A})\ \forall \ell$.*

orthogonal iteration without shift	*QR-iteration with shift*
$\widehat{\mathbf{Q}}_0 := I$	$\mathbf{A}_0 := \mathbf{A}$
$\left(\mathbf{A} - \mu^{(\ell)}\right) \widehat{\mathbf{Q}}_\ell =: \widehat{\mathbf{Q}}_{\ell+1}\mathbf{R}_{\ell+1}$	$\mathbf{A}_\ell - \mu^{(l)} =: \mathbf{Q}_{\ell+1}\mathbf{R}_{\ell+1}$
	$\mathbf{A}_{\ell+1} := \mathbf{R}_{\ell+1}\mathbf{Q}_{\ell+1} + \mu^{(\ell)}$

Then: $\forall \ell$:

(i) $\left(\mathbf{A} - \mu^{(\ell)}\right)\left(\mathbf{A} - \mu^{(\ell-1)}\right)\dots\left(\mathbf{A} - \mu^{(0)}\right)\widehat{\mathbf{Q}}_0 = \widehat{\mathbf{Q}}_{\ell+1}\widehat{\mathbf{R}}_{\ell+1}$ with $\widehat{\mathbf{R}}_{\ell+1} = \mathbf{R}_{\ell+1}\dots\mathbf{R}_1$

(ii) $\mathbf{A}_\ell = \widehat{\mathbf{Q}}_\ell^H \mathbf{A} \widehat{\mathbf{Q}}_\ell$

(iii) $\widehat{\mathbf{Q}}_\ell = \mathbf{Q}_1 \cdots \mathbf{Q}_\ell$

Proof: Exercise. Define the matrices $\mathbf{Q}_\ell$ and $\mathbf{R}_\ell$ by the QR-iteration with shift, i.e., by $\mathbf{A}_\ell - \mu^{(\ell)} = \mathbf{Q}_\ell\mathbf{R}_\ell$. *Define* the matrices $\widehat{\mathbf{Q}}_\ell := \mathbf{Q}_1 \cdots \mathbf{Q}_\ell$ and $\widehat{\mathbf{R}}_\ell := \mathbf{R}_\ell \cdots \mathbf{R}_1$. Then (iii) is satisfied by definition. To see (ii) we compute

$$\mathbf{A}_{\ell+1} = \mathbf{R}_{\ell+1}\mathbf{Q}_{\ell+1} + \mu^{(\ell+1)} = \mu^{(\ell+1)} + \mathbf{Q}_{\ell+1}^H(\mathbf{A}_\ell - \mu^{(\ell+1)})\mathbf{Q}_{\ell+1} = \mathbf{Q}_{\ell+1}^H \mathbf{A}_\ell \mathbf{Q}_{\ell+1}.$$

Hence, an induction argument will show (ii).
We now show that matrices $\widehat{\mathbf{Q}}_\ell$ defined above satisfy (i). To that end, we compute

$$\widehat{\mathbf{Q}}_{\ell+1}\widehat{\mathbf{R}}_{\ell+1} = \widehat{\mathbf{Q}}_\ell\mathbf{Q}_{\ell+1}\mathbf{R}_{\ell+1}\widehat{\mathbf{R}}_\ell = \mathbf{Q}_\ell(\mathbf{A}_\ell - \mu^{(\ell)})\widehat{\mathbf{R}}_\ell \overset{(ii)}{=} \widehat{\mathbf{Q}}_\ell(\widehat{\mathbf{Q}}_\ell^H \mathbf{A}\widehat{\mathbf{Q}}_\ell - \mu^{(\ell)})\widehat{\mathbf{R}}_\ell = (\mathbf{A} - \mu^{(\ell)})\widehat{\mathbf{Q}}_\ell\widehat{\mathbf{R}}_\ell \tag{7.8}$$

Hence, an induction argument shows (i).
It remains to see that the $\widehat{\mathbf{Q}}_\ell$ actually satisfy

$$(\mathbf{A} - \mu^{(\ell)})\widehat{\mathbf{Q}}_\ell = \mathbf{Q}_{\ell+1}\mathbf{R}_{\ell+1}.$$

Since the $\mathbf{R}_i$ are invertible, this follows from (7.8) by multiplying both sides with $\widehat{\mathbf{R}}_\ell^{-1}$, which gives

$$\widehat{\mathbf{Q}}_{\ell+1}\mathbf{R}_{\ell+1} = \widehat{\mathbf{Q}}_{\ell+1}\widehat{\mathbf{R}}_{\ell+1}\widehat{\mathbf{R}}_\ell^{-1} \overset{(7.8)}{=} (\mathbf{A} - \mu^{(\ell)})\widehat{\mathbf{Q}}_\ell\widehat{\mathbf{R}}_\ell\widehat{\mathbf{R}}_\ell^{-1} = (\mathbf{A} - \mu^{(\ell)})\widehat{\mathbf{Q}}_\ell.$$

□

We have observed in Remark 7.23 that the orthogonal iteration (with $\mathbf{X}_0 = I$) performs several orthogonal iterations simultaneously. That is, the first k columns of $\widehat{\mathbf{Q}}_\ell$ are an ONB of the space $\mathbf{A}^\ell[\mathbf{e}_1, \dots, \mathbf{e}_k]$. More generally, Lemma 7.33 shows that the first k columns of $\widehat{\mathbf{Q}}_\ell$ are an ONB of $(\mathbf{A} - \mu^{(\ell)}) \cdots (\mathbf{A} - \mu^{(0)})[\mathbf{e}_1, \dots, \mathbf{e}_k]$.
The following Lema 7.34 shows that in the case without shift that $(\mathbf{A}^H)^{-\ell}\mathbf{e}_n$ is a multiple of the last column of $\widehat{\mathbf{Q}}_\ell$:

Lemma 7.34 *Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be invertible. Define the permutation matrix*

$$\mathbf{P} = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}$$

$[\![\ \mathbf{BP} = \mathbf{B}(:, [n : -1 : 1]); \mathbf{P}^H\mathbf{B} = \mathbf{PB} = \mathbf{B}([n : -1 : 1], :)\ ]\!]$
Let $\mathbf{A}^\ell = \widehat{\mathbf{Q}}_\ell\widehat{\mathbf{R}}_\ell$ with $\widehat{\mathbf{Q}}_\ell$ unitary and $\widehat{\mathbf{R}}_\ell$ upper triangular. Then:

$$(\mathbf{A}^H)^{-\ell}\mathbf{e}_n = (\mathbf{A}^H)^{-\ell}\mathbf{P}\mathbf{e}_1 = \widehat{\mathbf{Q}}_\ell(\underbrace{\mathbf{P}\underbrace{(\mathbf{P}^H\widehat{\mathbf{R}}_\ell^{-H}\mathbf{P})\mathbf{e}_1}_{\|\mathbf{e}_1}}_{\|\mathbf{e}_n}) = \text{ multiple of last column of } \widehat{\mathbf{Q}}_\ell$$

Proof: direct calculation. □

Lemma 7.34 shows that the *last* columns of the matrices $\mathbf{Q}_\ell$ correspond to an inverse iteration for $\mathbf{A}^H$. More generally, one can show for the case with shifts:

Lemma 7.35

$$(\mathbf{A}^H - \overline{\mu^{(\ell)}})^{-1} \cdots (\mathbf{A}^H - \overline{\mu^{(0)}})^{-1}\mathbf{e}_n = \textit{multiple of } \widehat{\mathbf{Q}}_\ell(:,n).$$

with $\widehat{\mathbf{Q}}_\ell$ given by Lemma 7.33.

Proof: Computation/literature. □

That is, the last column of $\widehat{\mathbf{Q}}_\ell$ corresponds to an inverse iteration for $\mathbf{A}^H$ with shifts related to the shifts of the QR-iteration. Hence it is sensible to select the shifts $\mu^{(\ell)}$ of the QR-iteration such that $\overline{\mu^{(\ell)}}$ is the Rayleigh quotient for $\mathbf{q}_n := \widehat{\mathbf{Q}}_\ell(:,n)$:

$$\overline{\mu^{(\ell)}} := \frac{\mathbf{q}_n^H \mathbf{A}^H \mathbf{q}_n}{\|\mathbf{q}_n\|_2^2} = \mathbf{q}_n^H \mathbf{A}^H \mathbf{q}_n = (\widehat{\mathbf{Q}}_\ell \mathbf{e}_n)^H \mathbf{A}^H (\widehat{\mathbf{Q}}_\ell \mathbf{e}_n).$$

Hence,

$$\mu^{(\ell)} := \left((\widehat{\mathbf{Q}}_\ell \mathbf{e}_n)^H \mathbf{A}^H (\widehat{\mathbf{Q}}_\ell \mathbf{e}_n)\right)^H = \mathbf{e}_n^H \widehat{\mathbf{Q}}_\ell^H \mathbf{A} \widehat{\mathbf{Q}}_\ell \mathbf{e}_n = \mathbf{e}_n^H \mathbf{A}_{\ell+1} \mathbf{e}_n = \mathbf{A}_{\ell+1}(n,n).$$

That is, the shift should be taken as the bottom lower entry $\mathbf{A}_{\ell+1}(n,n)$ of $\mathbf{A}_{\ell+1}$.

Exercise 7.36 *Show that if μ is an eigenvalue of $\mathbf{A}$, then $\overline{\mu}$ is an eigenvalue of $\mathbf{A}^H$.* ∎

A few comments are in order:

1. The general behavior of the QR-algorithm with shift is that one has rapid convergence (quadratic convergence!) towards one eigenvalue since it behaves like a Rayleigh quotient method. Furthermore, one has linear convergence towards the remaining eigenvalues.

2. The rapid convergence towards one eigenvalue make deflation possible → iterate on a smaller matrix!

3. For deflation, monitor $\mathbf{A}_\ell(n-1,n)$: Since one will perform the QR-algorithm for $\mathbf{A}_0$ in Hessenberg form (so that all $\mathbf{A}_\ell$ have Hessenberg form — cf. Remark 7.25) $\mathbf{A}_\ell$ is Hessenberg, and it has only two non-zero entries in the nth row, namely, $\mathbf{A}_\ell(n,n-1)$ and $\mathbf{A}_\ell(n,n)$. Hence, deflation can be done when $\mathbf{A}_\ell(n,n-1)$ is sufficiently small (e.g., a small multiple of machine precision).

slide 35a

7.7.1 further comments on QR

Problem: in particular, for real matrices with eigenvalues appearing in complex conjugate pairs, it is possible for the Rayleigh quotient method to fail: $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then the QR-iteration (with shift) yields $\mathbf{A}_\ell = \mathbf{A} \forall \ell$.
solution:(Wilkinson-shift):
consider the two eigenvalues λ_1, λ_2 of $\mathbf{A}(n-1:n, n-1:n)$ and choose the shift as the eigenvalue that is closer to $\mathbf{A}(n,n)$.
Problem: QR-algorithm does not converge with Wilkinson shift:

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad \sigma(\mathbf{A}) = \{1, \frac{1}{2}(-1+\sqrt{3}\mathbf{i}), \frac{1}{2}(-1-\sqrt{3}\mathbf{i})\}$$

here, the (Wilkinson) shift is 0 and all eigenvalues have absolute value 1. Indeed, $\mathbf{A}_\ell = \mathbf{A}$ for all ℓ.
solution: If the QR-iteration does not converge, then make a "random shift". In general, this leads to a separation (in absolute value) of the eigenvalues and thus convergence: If $\lambda_3 \neq \lambda_1 \neq \lambda_2 \neq \lambda_3$, but $|\lambda_1| = |\lambda_2| = |\lambda_3|$, then $|\lambda_2 - \lambda| \neq |\lambda_1 - \lambda| \neq |\lambda_2 - \lambda| \neq |\lambda_3 - \lambda|$.

7.7.2 real matrices

Suppose $\mathbf{A}$ is real and one is not interested in complex shifts (e.g., because one wishes to stay with real arithmetic). In this case, eigenvalues appear in complex conjugate pairs λ, $\overline{\lambda}$. One can therefore makes two QR-steps with shifts λ and $\overline{\lambda}$. It is possible to combine these two steps purely in real arithmetic.