Utrecht, 30 november 2016

Eigenvalues and eigenvectors

Gerard Sleijpen

Universiteit Utrecht
Department of Mathematics

http://www.staff.science.uu.nl/~sleij101/

Subspace methods

Iterate until sufficiently accurate:

- **Expansion.** Expand the search subspace $\mathcal{V}_k$.
  Restart if $\dim(\mathcal{V}_k)$ is too large.
- **Extraction.** Extract an appropriate approximate solution from the search subspace.

Example. Krylov subspace methods
as GMRES, CG, Arnoldi, Lanczos: expansion by $\mathbf{t}_k = \mathbf{A}\mathbf{v}_k$

Goal.
Expansion. $\angle(\mathbf{x}, \mathcal{V}_{k+1}) \ll \angle(\mathbf{x}, \mathcal{V}_k)$
Extraction. Find $\mathbf{u} \in \mathcal{V}_k$ s.t. $\angle(\mathbf{x}, \mathbf{u}) \approx \angle(\mathbf{x}, \mathcal{V}_{k+1})$

Program Lecture 12

Extracting eigenpairs

- Extraction
- Ritz values and harmonic Ritz values

Perturbed eigenproblems

- Errors and perturbations
- Miscellenuous results
- Accuracy eigenvalues versus eigenvectors
- Perturbed eigenpairs
- Forward error and residual
- Pseudo spectra

Extraction strategies

Let $\mathcal{V} \equiv \text{span}(\mathbf{V})$ be a search subspace.

Find $\mathbf{u} \equiv \mathbf{V}y \in \mathcal{V}$ such that

- **(Ritz–)Galerkin.** $\mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{V}$ Ritz values
  Orthogonal residuals $\mathbf{Au} - \mathbf{b} \perp \mathbf{V}$ for solving $\mathbf{Ax} = \mathbf{b}$
- **Petrov–Galerkin.** $\mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{AV}$ harmonic Ritz values.
  Minimal residuals for solving $\mathbf{Ax} = \mathbf{b}$:
  $$\mathbf{u} = \text{minarg}_{\mathbf{z}}\|\mathbf{Az} - \mathbf{b}\|_2 \quad \Leftrightarrow \quad \mathbf{Au} - \mathbf{b} \perp \mathbf{AV}$$
- **Refined Ritz.** For a given approximate eigenvalue ϑ,
  $$\mathbf{u} \equiv \text{minarg}_{\tilde{\mathbf{u}} \in \mathcal{V}}\|\mathbf{A}\tilde{\mathbf{u}} - \vartheta\tilde{\mathbf{u}}\|_2$$

Selection

Ritz–Galerkin and Petrov–Galerkin lead to
k Ritz pairs $(\vartheta_i, \mathbf{u}_i)$, Petrov pairs, respectively $(i = 1, \dots, k)$.

Select the most 'promising' one as approximate eigenpair.

'**Most promising**':
1) Formulate a property that, among all eigenpairs, characterizes the wanted eigenpair
Example. $\lambda = \max(\mathrm{Re}(\lambda_j))$, $\lambda = \min|\lambda_j|$, $\lambda = \min|\lambda_j - \tau|$, $\dots$.

2) Select among all Ritz pairs the one with this property.
Example. $\vartheta = \max(\mathrm{Re}(\vartheta_i))$, $\vartheta = \min|\vartheta_i|$, $\vartheta = \min|\vartheta_i - \tau|$, $\dots$.

Warning. May lead to a 'wrong' selection

One wrong selection = one 'useless' iteration step.

One wrong selection at restart may spoil convergence.

Ritz values

Proposition. $\mathbf{u} = \mathbf{V}y$. Ritz values are Rayleigh quotients:

$$\mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{V} \quad \Rightarrow \quad \vartheta = \rho(\mathbf{u}) \equiv \frac{\mathbf{u}^*\mathbf{Au}}{\mathbf{u}^*\mathbf{u}}.$$

Proposition. For a given approximate eigenvector $\mathbf{u}$, the Rayleigh quotient is best approximate eigenvalue, i.e., gives the smallest residual:

$$\|\mathbf{Au} - \vartheta\mathbf{u}\|_2 \leq \|\mathbf{Au} - \tilde{\vartheta}\mathbf{u}\|_2 \quad (\tilde{\vartheta} \in \mathbb{C}) \quad \Rightarrow \quad \vartheta = \rho(\mathbf{u}).$$

Proof.

$$\mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{V} \;\Rightarrow\; \mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{V}y = \mathbf{u} \;\Leftrightarrow\; \vartheta = \rho(\mathbf{u}).$$

$$\|\mathbf{Au} - \vartheta\mathbf{u}\|_2 \leq \|\mathbf{Au} - \tilde{\vartheta}\mathbf{u}\|_2 \; (\tilde{\vartheta} \in \mathbb{C}) \;\Leftrightarrow\; \mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{u}.$$

Ritz values

For ease of discussion,

assume $\mathbf{AX} = \mathbf{X}\Lambda$ with $\mathbf{X}^*\mathbf{X} = \mathbf{I}$,

where $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$:

- $\mathbf{Ax}_i = \lambda_i \mathbf{x}_i \qquad (i = 1, \dots, n)$,
- the eigenvectors $\mathbf{x}_i$ form an orthonormal basis of $\mathbb{C}^n$.

Terminology. **A** has an orthonormal basis **X** of eigenvectors.

Note. **A** is **normal** iff $\mathbf{A}^*\mathbf{A} = \mathbf{AA}^*$.

Hermitian and unitary matrices are normal.

A is normal $\Leftrightarrow$
A has an orthonormal basis of eigenvectors.

Ritz values

For ease of discussion,

assume $\mathbf{AX} = \mathbf{X}\Lambda$ with $\mathbf{X}^*\mathbf{X} = \mathbf{I}$.

$\mathbf{u}$ approximate eigenvector, $\|\mathbf{u}\|_2 = 1$, $\vartheta = \rho(\mathbf{u})$.

$$\mathbf{u} = \sum \beta_i \mathbf{x}_i \;\text{ with }\; \textstyle\sum_i |\beta_i|^2 = 1,$$

$$\vartheta = \rho(\mathbf{u}) = \textstyle\sum_i |\beta_i|^2 \lambda_i.$$

Proposition. If **A** is normal, then any Ritz value is a convex mean (i.e., weighted averages) of eigenvalues.

Proposition. Ritz values form

- a safe selection for finding extremal eigenvalues,
- an unsafe selection for interior eigenvalues.

Harmonic Ritz values

For ease of discussion,

assume $\quad \mathbf{AX} = \mathbf{X}\Lambda \quad$ with $\quad \mathbf{X}^*\mathbf{X} = \mathbf{I}$.

Assume • we are interested in eigenvalue λ closest to 0,
• 0 is in the interior of the spectrum, • $\lambda \neq 0$.

Note that $\mathbf{A}^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$ and $\frac{1}{\lambda}$ extremal in $\{\frac{1}{\lambda_i}\}$

With respect to $\mathbf{W}$, find $\widetilde{\mathbf{x}} \equiv \mathbf{W}y$ st $\mathbf{A}^{-1}\widetilde{\mathbf{x}} - \mu\widetilde{\mathbf{x}} \perp \mathbf{W}$:

largest μ forms a safe selection ($\Rightarrow \lambda \approx \frac{1}{\mu}$, $\widetilde{\mathbf{x}} \approx \mathbf{x}$)

Select $\mathbf{W} = \mathbf{AV}$. Then, with $\mathbf{u} \equiv \mathbf{V}y$, we have $\widetilde{\mathbf{x}} = \mathbf{Au}$

$$\mathbf{A}^{-1}\widetilde{\mathbf{x}} - \mu\widetilde{\mathbf{x}} \perp \mathbf{W} \qquad \Leftrightarrow \qquad \frac{1}{\mu}\mathbf{u} - \mathbf{Au} \perp \mathbf{AV}$$

Proposition. Harmonic Ritz values form a safe selection for finding eigenvalues in the interior (close to 0).

Harmonic Ritz values

For ease of discussion,

assume $\quad \mathbf{AX} = \mathbf{X}\Lambda \quad$ with $\quad \mathbf{X}^*\mathbf{X} = \mathbf{I}$.

Assume • we are interested in eigenvalue λ closest to 0,
• 0 is in the interior of the spectrum, • $\lambda \neq 0$.

Strategy using harmonic Ritz values

1) Solve $\quad \mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{AV}$
2) Select ϑ closest to 0.

Proposition. If $\mathbf{A}$ is normal, then harmonic Ritz values are **harmonic** means of the eigenvalues.

$\mathbf{A}$ given $n \times n$ matrix, $\quad \mathbf{Ax} = \lambda\mathbf{x}$

In practice: Only *approximate* eigenpairs $(\vartheta, \mathbf{u})$ can be computed, $\vartheta \in \mathbb{C}$, $\mathbf{u}$ a non-trivial n-vector.

$\left\{\begin{array}{ll} \lambda - \vartheta & \textbf{forward error} \text{ in the appr. eigenvalue} \\ \angle(\mathbf{x}, \mathbf{u}) & \textbf{forward error} \text{ in the appr. eigenvector} \end{array}\right.$

with **residual** $\mathbf{r} \equiv \mathbf{Au} - \vartheta\mathbf{u}$.

A perturbation Δ of $\mathbf{A}$ such that

$$(\mathbf{A} - \Delta)\mathbf{u} = \vartheta\mathbf{u}$$

is called a **backward error** of the appr. eigenpair.

Proposition. With $\|\mathbf{u}\|_2 = 1$ and $\Delta \equiv \mathbf{r}\mathbf{u}^*$, we have

$$(\mathbf{A} - \Delta)\mathbf{u} = \vartheta\mathbf{u} \quad \& \quad \|\Delta\|_2 \leq \|\mathbf{r}\|_2$$

Backward error

$(\vartheta, \mathbf{u})$ with $\vartheta \in \mathbb{C}$, $\mathbf{u}$ a non-trivial n-vector is an approximate eigenpair if the **residual** $\mathbf{r} \equiv \mathbf{Au} - \vartheta\mathbf{u}$ is small.

Proposition. With $\|\mathbf{u}\|_2 = 1$ and $\Delta \equiv \mathbf{r}\mathbf{u}^*$, we have

$$(\mathbf{A} - \Delta)\mathbf{u} = \vartheta\mathbf{u} \quad \& \quad \|\Delta\|_2 \leq \|\mathbf{r}\|_2$$

For a given approximate eigenvector $\mathbf{u}$,
we have the smallest residual

$$\vartheta = \text{argmin}_\mu \|\mathbf{Au} - \mu\mathbf{u}\|_2 \quad \Leftrightarrow \quad \mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{u} \quad \Leftrightarrow \quad \vartheta = \frac{\mathbf{u}^*\mathbf{Au}}{\mathbf{u}^*\mathbf{u}}$$

$\rho(\mathbf{u}) \equiv \frac{\mathbf{u}^*\mathbf{Au}}{\mathbf{u}^*\mathbf{u}}$ is the **Rayleigh quotient** (of $\mathbf{u}$ wrt $\mathbf{A}$).

Note. If ϑ is the Rayleigh quotient, then $\mathbf{r} \perp \mathbf{u}$.

Backward error

$(\vartheta, \mathbf{u})$ with $\vartheta \in \mathbb{C}$, $\mathbf{u}$ a non-trivial n-vector is an approximate eigenpair if the **residual** $\mathbf{r} \equiv \mathbf{A}\mathbf{u} - \vartheta\mathbf{u}$ is small.

Proposition. With $\|\mathbf{u}\|_2 = 1$ and $\Delta \equiv \mathbf{r}\mathbf{u}^*$, we have

$$(\mathbf{A} - \Delta)\mathbf{u} = \vartheta\mathbf{u} \quad \& \quad \|\Delta\|_2 \le \|\mathbf{r}\|_2$$

- How do eigenpairs respond to perturbations?
- How to find (approximate) eigenpairs (with small residuals).

Note. Δ may be structured.

Here, we will pay special attention only to $\Delta = \mathbf{r}\mathbf{u}^*$, i.e., structure from backward error.

Useful results, $\mathbf{A}^* = \mathbf{A}$

Theorem [Courant–Fischer] If $\lambda_1 \le \ldots \le \lambda_n$, then

$$\lambda_i = \min_{\mathcal{W}} \max_{\mathbf{w}} \rho(\mathbf{w}) \qquad (i = 1, \ldots, n),$$

where the maximum is taken over all non-zero $\mathbf{w} \in \mathcal{W}$ and the minimum over all i-dimensional subspaces $\mathcal{W}$.

Theorem [Cauchy interlace] The eigenvalues of $\mathbf{A}$,

if $\mathbf{A} = \begin{bmatrix} \mathbf{H} & \mathbf{b} \\ \mathbf{b}^* & \alpha \end{bmatrix}$, and $\mathbf{H}$ **interlace**:

$$\lambda_1(\mathbf{A}) \le \lambda_1(\mathbf{H}) \le \lambda_2(\mathbf{A}) \le \lambda_2(\mathbf{H}) \le \ldots \le \lambda_{n-1}(\mathbf{H}) \le \lambda_n(\mathbf{A})$$

Useful result for Hermitian problems using subspace methods, where, per step, the projected matrix is extended with one row and one column.

Accuracy eigenvalues versus eigenvectors

The approximate eigenvalue is usually much more accurate then the eigenvector.

If $\mathbf{A}$ is Hermitian, then the error in the eigenvalue is of order square of the error of the eigenvector.

Let $\mathbf{A}$ be Hermitian: $\mathbf{A}^* = \mathbf{A}$.

Theorem. $|\rho(\mathbf{u}) - \lambda| \le \sin^2 \angle(\mathbf{x}, \mathbf{u}) \cdot \max_i |\lambda_i - \lambda|$.

Theorem. If $\lambda = \lambda_1 < \lambda_i$ all $i > 1$, then

$$\sin^2 \angle(\mathbf{x}_1, \mathbf{u}) \le \frac{\rho(\mathbf{u}) - \lambda_1}{\lambda_2 - \lambda_1}.$$

Proofs. Write $\mathbf{u} = c\mathbf{x} + s\mathbf{z}$, where $\mathbf{z} \perp \mathbf{x}$ and $\|\mathbf{z}\|_2 = 1$.

$$\rho(\mathbf{u}) - \lambda = \mathbf{u}^*(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = s^2\mathbf{z}^*(\mathbf{A} - \lambda)\mathbf{z}$$

and $\rho(\mathbf{z}) = \mathbf{z}^*\mathbf{A}\mathbf{z}$ is in the convex hull of $\{\lambda_j \mid j \ne j_0\}$.
In case $\mathbf{x} = \mathbf{x}_1$ we have that $\rho(\mathbf{z}) \ge \lambda_2$ (Courant–Fischer).

Examples

$$\begin{bmatrix} 0 & \tau \\ \tau & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ \tau & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & \tau \\ \tau & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & \tau \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \tau & 1 \\ 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & \tau \\ \tau & 0 & 0 \end{bmatrix}$$

If $\lambda = \lambda_i(0)$ is a non-simple eigenvalue of $\mathbf{A}(\tau)$ at $\tau = 0$, then $\lambda_i(\tau)$ can be expressed as a **Puiseux** series

$$\lambda_{i+m}(\tau) = \lambda + \sum_{j=1}^{\infty} \alpha_j\, \omega^m\, \eta^j, \quad (m = 0, 1, \ldots, p-1, \tau \approx 0),$$

where $\eta \equiv r^{1/p} e^{i\phi/p}$ if $\tau = r\, e^{i\phi}$, and $\omega \equiv e^{2\pi i/p}$, $p \le \text{mult}(\lambda)$.

Smooth perturbations

For $\tau \in \mathbb{C}$, consider $\mathbf{A}(\tau) \equiv \mathbf{A} - \tau\mathbf{E}$.

Then $\mathbf{A}(0) = \mathbf{A}$ and $\tau \rightsquigarrow \mathbf{A}(\tau)$ is smooth.

Theorem.

- There are continuous functions $\tau \rightsquigarrow \lambda_j(\tau)$ such that

$$\lambda_1(\tau), \ldots, \lambda_n(\tau) \text{ are the eigenvalues of } \mathbf{A}(\tau)$$

counted according to multiplicity $(\tau \in \mathbb{C})$.

- If $\lambda_j(0)$ is a simple eigenvalue of $\mathbf{A}(0)$, then

$$\tau \rightsquigarrow \lambda_j(\tau) \text{ is analytic for } \tau \approx 0.$$

If, for some vector $\mathbf{w}$, the associated eigenvector $\mathbf{x}_j(\tau)$ is scaled st $\mathbf{w}^*\mathbf{x}_j(\tau) = 1$, then $\tau \rightsquigarrow \mathbf{x}_j(\tau)$ is also analytic.

- If $\mathbf{A}(\tau)$ is Hermitian $(\tau \in \mathbb{R})$, then there are eigenvalues $\lambda_j(\tau)$ and eigenvectors $\mathbf{x}_j(\tau)$ that depend analytically on τ $(j = 1, \ldots, n), \tau \approx 0$.

Analysis strategy

To avoid technical details, we focuss on **simple** eigenvalues:
$\lambda = \lambda(0)$ is an eigenvalue of $\mathbf{A} = \mathbf{A}(0)$ of multiplicity 1.
$\mathbf{x} = \mathbf{x}(0)$ is the associated normalised eigenvector.

We will identify convenient non-singular matrices $\mathbf{V}$ (i.e., basis transforms) such that

$$\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \begin{bmatrix} \lambda & \mathbf{a}^* \\ \mathbf{0} & \mathbf{A}_1 \end{bmatrix} \quad \text{and} \quad \mathbf{V}^{-1}\mathbf{E}\mathbf{V} = \begin{bmatrix} \nu & \mathbf{f}^* \\ \tilde{\mathbf{r}} & \mathbf{E}_1 \end{bmatrix}$$

Special cases:
- $\epsilon\,\mathbf{E} = \Delta$ and $\epsilon \ll 1$
- $\epsilon\,\mathbf{E} = \mathbf{r}\mathbf{u}^*$ (rank 1)
- $\mathbf{A}$ normal $(\mathbf{V}^*\mathbf{V} = \mathbf{I})$
- $\mathbf{A}$ and $\mathbf{E}$ Hermitian $(\mathbf{V} = \mathbf{X}$ and $\mathbf{V}^{-1}\mathbf{A}\mathbf{V}$ diagonal$)$
- Combinations

The conditioning of an eigenvector

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad \lambda \text{ simple}, \quad \|\mathbf{x}\|_2 = 1$$

With $$\widetilde{\mathbf{A}} \equiv (\mathbf{I} - \mathbf{x}\mathbf{x}^*)\mathbf{A}(\mathbf{I} - \mathbf{x}\mathbf{x}^*)$$

and taking the inverse of $\widetilde{\mathbf{A}} - \lambda\mathbf{I}$ on $\mathbf{x}^\perp$, we have

Theorem. For some $(\vartheta, \mathbf{u})$ with $(\mathbf{A} - \Delta)\mathbf{u} = \vartheta\,\mathbf{u}$, we have

$$\tan \angle(\mathbf{x}, \mathbf{u}) \lesssim \|(\widetilde{\mathbf{A}} - \lambda\mathbf{I})^{-1}\|_2\,\|\Delta\|_2$$

$$\boxed{\text{Cond}_{\mathbf{x}}(\mathbf{A}) \equiv \|(\widetilde{\mathbf{A}} - \lambda\mathbf{I})^{-1}\|_2}$$

Interpretation. $\mathbf{x}_1, \ldots, \mathbf{x}_n$ orthonormal (i.e., $\mathbf{A}$ normal) $\Rightarrow$

$$\|(\widetilde{\mathbf{A}} - \lambda\mathbf{I})^{-1}\|_2 = \max\left\{\frac{1}{|\lambda_j - \lambda|} \mid \lambda_j \neq \lambda\right\} = \frac{1}{\gamma}$$

$\gamma \equiv \min_{\lambda_j \neq \lambda} |\lambda_j - \lambda|$ is the **spectral gap** for λ.

The conditioning of an eigenvalue

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{y}^*\mathbf{A} = \lambda\mathbf{y}^*, \quad \lambda \text{ simple}, \quad \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$$

Theorem. For some $(\vartheta, \mathbf{u})$ with $(\mathbf{A} - \Delta)\,\mathbf{u} = \vartheta\,\mathbf{u}$, we have

$$|\vartheta - \lambda| = \frac{|\mathbf{y}^*\Delta\,\mathbf{x}|}{|\mathbf{y}^*\mathbf{x}|} + \mathcal{O}(\|\Delta\|_2^2) \lesssim \frac{\|\Delta\|_2}{|\mathbf{y}^*\mathbf{x}|}$$

$$\boxed{\text{Cond}_\lambda(\mathbf{A}) \equiv \frac{1}{\cos \angle(\mathbf{x}, \mathbf{y})}}$$

Theorem [Weyl] If $\mathbf{A} = \mathbf{A}^*$ and $\Delta = \Delta^*$, then

$$|\lambda_i(\mathbf{A} + \Delta) - \lambda_i(\mathbf{A})| \leq \|\Delta\|_2.$$

$(\mathbf{y}^*\mathbf{x} = 1$, $\mathcal{O}(\tau^2)$-term is 0.) In this case, we even have

$$\lambda_1(\Delta) \leq \lambda_i(\mathbf{A} + \Delta) - \lambda_i(\mathbf{A}) \leq \lambda_n(\Delta)$$

Proof. Apply Courant–Fischer.

The conditioning of an eigenvalue

$\mathbf{Ax} = \lambda\mathbf{x}, \quad \mathbf{y}^*\mathbf{A} = \lambda\mathbf{y}^*, \quad \lambda$ simple, $\quad \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

Theorem. For some $(\vartheta, \mathbf{u})$ with $(\mathbf{A}-\Delta)\,\mathbf{u} = \vartheta\,\mathbf{u}$, we have

$$|\vartheta - \lambda| = \frac{|\mathbf{y}^*\Delta\,\mathbf{x}|}{|\mathbf{y}^*\mathbf{x}|} + \mathcal{O}(\|\Delta\|_2^2) \lesssim \frac{\|\Delta\|_2}{|\mathbf{y}^*\mathbf{x}|}$$

$$\boxed{\mathrm{Cond}_\lambda(\mathbf{A}) \equiv \frac{1}{\cos\angle(\mathbf{x},\mathbf{y})}}$$

$\Delta = \mathbf{r}\mathbf{u}^*$. **Theorem [Bauer-Fike].** If $\mathbf{A}$ is normal, then

$$|\vartheta - \lambda| \leq \|\mathbf{r}\|_2 \quad \text{for some} \quad \lambda \in \Lambda(\mathbf{A}).$$

$\Delta = \mathbf{r}\mathbf{u}^*$. **Theorem.** If $\mathbf{A}$ is normal and $\vartheta = \rho(\mathbf{u})$, then

$$\|\mathbf{r}\|_2 \leq \tfrac{1}{2}\gamma \quad \Rightarrow \quad |\rho(\mathbf{u}) - \lambda| \leq \frac{\|\mathbf{r}\|_2^2}{\gamma - \|\mathbf{r}\|}$$

(with γ the spectral gap for λ; $\mathbf{y}^*\Delta\mathbf{x} = \mathbf{u}^*\mathbf{r}\mathbf{u}^*\mathbf{u} = 0$).

Quantifying perturbations

$$\left(\begin{bmatrix} \lambda & \mathbf{a}^* \\ \mathbf{0} & \mathbf{A}_1 \end{bmatrix} - \tau \begin{bmatrix} \nu & \mathbf{f}^* \\ \widetilde{\mathbf{r}} & \mathbf{E}_1 \end{bmatrix}\right) \begin{bmatrix} 1 \\ \mathbf{z}_\tau \end{bmatrix} = \lambda(\tau) \begin{bmatrix} 1 \\ \mathbf{z}_\tau \end{bmatrix}, \qquad (*)$$

with $\quad \lambda(0) = \lambda$ and $\mathbf{z}_0 = \mathbf{0}$.

In our application, $\tau = \epsilon$, $\epsilon\mathbf{E} = \Delta$, and we can take

- $\mathbf{V} = \mathbf{X}$, the basis of eigenvectors $\rightsquigarrow$ Bauer–Fike,
- $\mathbf{V} = [\mathbf{x}, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ with $(\mathbf{v}_2, \ldots, \mathbf{v}_n)$ orthonormal basis $\mathbf{x}^\perp$.
- $\mathbf{V} = [\mathbf{x}, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ with $(\mathbf{v}_2, \ldots, \mathbf{v}_n)$ orthonormal basis $\mathbf{y}^\perp$.
  Here $\mathbf{y}$ is the normalised left eigenvector for λ.

Estimates based on the asymptotic expression from the preceding transparencies have to be multiplied by $\mathcal{C}_2(\mathbf{V})$.

Quantifying perturbations

$$\left(\begin{bmatrix} \lambda & \mathbf{a}^* \\ \mathbf{0} & \mathbf{A}_1 \end{bmatrix} - \tau \begin{bmatrix} \nu & \mathbf{f}^* \\ \widetilde{\mathbf{r}} & \mathbf{E}_1 \end{bmatrix}\right) \begin{bmatrix} 1 \\ \mathbf{z}_\tau \end{bmatrix} = \lambda(\tau) \begin{bmatrix} 1 \\ \mathbf{z}_\tau \end{bmatrix}, \qquad (*)$$

with $\quad \lambda(0) = \lambda$ and $\mathbf{z}_0 = \mathbf{0}$.

Note. $\|\mathbf{z}_\tau\|_2$ is the tangent of the angle between the eigenvector $(1, \mathbf{0}^\mathsf{T})^\mathsf{T}$ and the perturbed eigenvector $(1, \mathbf{z}_\tau^\mathsf{T})^\mathsf{T}$.

$$\begin{cases} \lambda - \tau\nu + \mathbf{a}^*\mathbf{z}_\tau - \tau\mathbf{f}^*\mathbf{z}_\tau = \lambda(\tau) \\ (\mathbf{A}_1 - \tau\mathbf{E}_1 - \lambda(\tau)\mathbf{I})\mathbf{z}_\tau = \tau\widetilde{\mathbf{r}} \end{cases}$$

Hence, for $\tau \to 0$,

$$\begin{cases} \mathbf{z}_\tau = \tau\,(\mathbf{A}_1 - \lambda\mathbf{I})^{-1}\widetilde{\mathbf{r}} + \mathcal{O}(\tau^2) \\ \lambda - \lambda(\tau) = \tau[\nu - \mathbf{a}^*(\mathbf{A}_1 - \lambda\mathbf{I})^{-1}\widetilde{\mathbf{r}}] + \mathcal{O}(\tau^2) \end{cases}$$

If $\mathbf{a} = \mathbf{0}$, then

$$\lambda - \lambda(\tau) = \tau\nu + \tau^2\mathbf{f}^*(\mathbf{A}_1 - \lambda\mathbf{I})^{-1}\widetilde{\mathbf{r}} + \mathcal{O}(\tau^3)$$

Perturbed eigenvalues and pseudo-spectra

$\mathcal{C}_2(\mathbf{X})$ is an bound for the conditioning of the eigenvalues.
However,
- it assumes a basis of eigenvectors,
- it does not discriminate between well conditioned and ill conditioned eigenvalues,
- it usually is not feasible to compute $\mathcal{C}_2(\mathbf{X})$.

The condition number $1/\cos\angle(\mathbf{y},\mathbf{x})$ of a simple eigenvalue depends on the angle between its left and right eigenvector. This number can be (accurately) computed for one or for a few eigenvalues.
However, in general it is not feasible to compute these numbers for all eigenvalues (for non-normal $\mathbf{A}$).
Moreover, for n large, the collection of all these numbers is too large to provide global information on the sensitivity of all eigenvalues to perturbations.

The **pseudo-spectrum** offers a graphical way to access the sensitivity of eigenvalues to perturbations. It gives information on individual eigenvalues, regardless multiplicity.

Perturbed eigenvalues and pseudo-spectra

For $\epsilon \geq 0$, the ϵ-**pseudo-spectrum** $\Lambda_\epsilon(\mathbf{A})$ is

$$\Lambda_\epsilon(\mathbf{A}) \equiv \bigcup \{\Lambda(\mathbf{A} + \Delta) \mid \|\Delta\|_2 \leq \epsilon\}$$

Proposition. $\vartheta \in \Lambda_\epsilon(\mathbf{A}) \ \Leftrightarrow$ smallest singular value $\mathbf{A} - \vartheta\mathbf{I} \leq \epsilon$

$$\Leftrightarrow \|(\mathbf{A} - \vartheta\mathbf{I})^{-1}\|_2^{-1} \leq \epsilon.$$

Observations.

- If $\lambda \in \Lambda(\mathbf{A})$ and $|\lambda - \vartheta| \leq \epsilon$, then $\vartheta \in \Lambda_\epsilon(\mathbf{A})$.
- Often the pseudo-spectrum is much bigger than the union of discs with radius ϵ around eigenvalues.
- Often the value of ϵ does not seem to play a significant role (reason: $\epsilon^{\frac{1}{32}} \approx 1$ for any $\epsilon \in [10^{-8}, 10^{+8}]$).
- In floating point arithmetic $\mathbf{c} \equiv \mathbf{Au}$ is exactly $\mathbf{c} = (\mathbf{A} + \Delta)\mathbf{u}$ for some small perturbation Δ.
- If $\mathbf{r} = \mathbf{Au} - \vartheta\mathbf{u} \ \Rightarrow \ \vartheta \in \Lambda_\epsilon(\mathbf{A})$ for $\epsilon \geq \|\mathbf{r}\|_2$.