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. $\angle(\mathbf{x}, \mathbf{u}) \approx \angle(\mathbf{x}, \mathcal{V}_{k+1})$

Krylov subspace methods

+ Polynomial approximation theory is applicable
+ Structure can be exploited (more efficient steps)
− Sensitive to errors
− Not flexible

Subspace methods

− (Slightly) more costly per step
+ Less sensitive to errors
+ Flexible (allowing faster convergence)

Subspace methods

We have introduced (Krylov) subspace methods as a technique to accelerate convergence of simple iteration methods (as power method).

However, the approach has additional advantages:

- It allows to (somewhat) steer convergence towards the wanted eigenpair(s), if, for instance, restarts are required.

- If more than one eigenpair is required, then the subspace upon convergence of the first eigenpair will contain good approximation to the second eigenpair, etc..

- If an eigenvalue is not simple, an eigenspace has to be computed rather than an eigenvector.

Subspace methods

Block methods also rely on subspaces, that is on matrices $\mathbf{V}_k$ that span a search subspace. However, in contrast to subspace methods, block methods keep the size of the matrix $\mathbf{V}_k$ fixed in each step k to $n \times \ell$. Block methods are actually "simple methods" aiming for solving the "block" eigenvalue problem

$$\mathbf{A}\mathbf{X}_\ell = \mathbf{X}_\ell \Lambda_\ell,$$

where $\mathbf{X}_\ell$ is an $n \times \ell$ matrix with eigenvectors as columns and Λ_ℓ is an $\ell \times \ell$ diagonal matrix with eigenvalues on its diagonal.

Nevertheless, block methods somewhat share the advantages of subspace methods over "simple iteration". Both approaches can be combined.

The QR-algorithm can be viewed as a block method.

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 $(\vartheta, \mathbf{u})$ from the search subspace.

This lecture, for $\mathbf{Ax} = \lambda\mathbf{x}$ with $\mathbf{A}$ $n \times n$, focusses on
extraction
sufficiently accurate
(residuals, errors, perturbations, etc.)

Notation.
$\mathbf{u}$ is an approximation of $\mathbf{x}$ (in direction) with
ϑ the approximate eigenvalue,
$\mathbf{r} \equiv \mathbf{Au} - \vartheta\mathbf{u}$ is the residual,

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

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,

  $$\mathbf{u} \equiv \text{minarg}_{\widetilde{\mathbf{u}} \in \mathcal{V}} \|\mathbf{A}\widetilde{\mathbf{u}} - \vartheta\widetilde{\mathbf{u}}\|_2$$

Selection

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

Select the most 'promising' one as approximate eigenpair.

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

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

Selection

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

Select the most ‘promising’ one as approximate eigenpair.

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

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

Warning. May lead to a ‘wrong’ selection:

there may be a more accurate Ritz pair than the selected one

Selection

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

Select the most 'promising' one as approximate eigenpair.

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

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

Warning. May lead to a 'wrong' selection.

One wrong selection = one 'useless' iteration step

Selection

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

Select the most ‘promising’ one as approximate eigenpair.

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

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

Warning. May lead to a ‘wrong’ selection

One wrong selection = one ‘useless’ iteration step.

One wrong selection at restart may spoil convergence.

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

Ritz values

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

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

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

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

Proof.

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

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

Ritz values

For ease of discussion,

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

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

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

Terminology. $\mathbf{A}$ has an orthonormal basis $\mathbf{X}$ of eigenvectors.

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

Hermitian and unitary matrices are normal.

$\mathbf{A}$ is normal $\Leftrightarrow$

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

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

because $(\mathbf{x}_i)$ orthonormal, $1 = \|\mathbf{u}\|_2 = \sum_i |\beta_i|^2$

Ritz values

For ease of discussion,

$$\text{assume} \qquad \mathbf{A}\mathbf{X} = \mathbf{X}\Lambda \qquad \text{with} \qquad \mathbf{X}^*\mathbf{X} = \mathbf{I}.$$

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

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

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

Ritz values

For ease of discussion,

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

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

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

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

Proposition. If $\mathbf{A}$ is normal, then any Ritz value is a convex mean (i.e., weighted averages) of eigenvalues.

Ritz values

For ease of discussion,

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

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

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

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

Proposition. If $\mathbf{A}$ is normal, then any Ritz value is a convex mean (i.e., weighted averages) of eigenvalues.

Assume $\vartheta \approx \lambda = \lambda_{j_0}$. Can we conclude that $\mathbf{u} \approx \mathbf{x} \equiv \mathbf{x}_{j_0}$?

Ritz values

For ease of discussion,

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

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

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

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

Proposition. If $\mathbf{A}$ is normal, then any Ritz value is a convex mean (i.e., weighted averages) of eigenvalues.

Assume $\vartheta \approx \lambda = \lambda_{j_0}$. Can we conclude that $\mathbf{u} \approx \mathbf{x} \equiv \mathbf{x}_{j_0}$?

- If λ extremal: $\vartheta \approx \lambda \quad \Leftrightarrow \quad |\beta_{j_0}| \approx 1$ & $|\beta_i|^2 \approx 0$ $(i \neq j_0)$

λ is extremal if it is a vertex of the convex hull of the spectrum of $\mathbf{A}$.

Ritz values

For ease of discussion,

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

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

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

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

Proposition. If $\mathbf{A}$ is normal, then any Ritz value is a convex mean (i.e., weighted averages) of eigenvalues.

Assume $\vartheta \approx \lambda = \lambda_{j_0}$. Can we conclude that $\mathbf{u} \approx \mathbf{x} \equiv \mathbf{x}_{j_0}$?

- If λ extremal: $\vartheta \approx \lambda \quad \Leftrightarrow \quad \mathbf{u} \approx \text{sign}(\beta_{j_0})\mathbf{x}$

Ritz values

For ease of discussion,

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

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

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

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

Proposition. If $\mathbf{A}$ is normal, then any Ritz value is a convex mean (i.e., weighted averages) of eigenvalues.

Assume $\vartheta \approx \lambda = \lambda_{j_0}$. Can we conclude that $\mathbf{u} \approx \mathbf{x} \equiv \mathbf{x}_{j_0}$?

- If λ extremal: $\vartheta \approx \lambda \quad \Leftrightarrow \quad \mathbf{u} \approx \text{sign}(\beta_{j_0})\mathbf{x}$
- If λ in the interior: $\vartheta \approx \lambda \quad \Leftrightarrow \quad$??? $\quad$ (**Ex.**)

Ritz values

For ease of discussion,

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

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

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

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

Proposition. If $\mathbf{A}$ 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 $\quad \mathbf{A}\mathbf{X} = \mathbf{X}\Lambda \quad$ 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, • $\lambda \neq 0$.

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 $\{\frac{1}{\lambda_i}\}$

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 ($\Rightarrow \lambda \approx \frac{1}{\mu}$, $\widetilde{\mathbf{x}} \approx \mathbf{x}$)

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 ($\Rightarrow \lambda \approx \frac{1}{\mu}$, $\widetilde{\mathbf{x}} \approx \mathbf{x}$)

How to select $\mathbf{W}$? How to avoid $\mathbf{A}^{-1}$?

Harmonic Ritz values

For ease of discussion,

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

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 $\widetilde{\mathbf{x}} = \mathbf{Au}$

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

Harmonic Ritz values

For ease of discussion,

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

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 $\widetilde{\mathbf{x}} = \mathbf{Au}$

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

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

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 $\mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{AV}$

2) Select ϑ closest 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.

Naming. $\mathbf{u} = \sum \beta_i \mathbf{x}_i, \quad \mathbf{Au} - \vartheta\mathbf{u} \perp \mathbf{Au} \;\Rightarrow$ [See Exercise 13.4]

$$\vartheta = \frac{(\mathbf{Au})^*(\mathbf{Au})}{(\mathbf{Au})^*\mathbf{u}} = \frac{\sum |\beta_i|^2\,|\lambda_i|^2}{\sum |\beta_i|^2\bar{\lambda}_i} = \frac{1}{\sum \alpha_i \frac{1}{\lambda_i}},$$

where $\alpha_i \equiv |\beta_i|^2\,|\lambda_i|^2/(\sum_i |\beta_i|^2\,|\lambda_i|^2)$, $\sum \alpha_i = 1$, $\alpha_i \geq 0$.

Harmonic Ritz values

For ease of discussion,

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

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

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

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 τ,
• τ is in the interior of the spectrum, • $\lambda \neq \tau$.

Strategy using harmonic Ritz values

1) Solve $\quad \mathbf{Au}-\vartheta\mathbf{u} \perp (\mathbf{A}-\tau\mathbf{I})\mathbf{V}$

2) Select ϑ closest to τ.

Refined Ritz

For a given approximate eigenvalue ϑ,
the **refined Ritz vector** $\mathbf{u}$ is such that

$$\mathbf{u} = \text{argmin}\{\|\mathbf{A}\widetilde{\mathbf{u}} - \vartheta\widetilde{\mathbf{u}}\|_2 \mid \widetilde{\mathbf{u}} \in \mathcal{V}\}$$

Refined Ritz is only useful if
ϑ is very close to the wanted eigenvalue.

A few extra digits of accuracy can be obtained
upon convergence.

Observation. If $\mathbf{A}^* = \mathbf{A}$, then a refined Ritz vector is a Ritz vector for $(\mathbf{A} - \vartheta\,\mathbf{I})^2$. Note that squaring also turns interior eigenvalues to the exterior.

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

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

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

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

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

The error can not be computed.

Alternative: compute the **residual** $\mathbf{r} \equiv \mathbf{A}\mathbf{u} - \vartheta\mathbf{u}$.

We will learn that,

If the scaled residual norm $\|\mathbf{r}\|_2 / \|\mathbf{u}\|_2$ is small,
then $(\vartheta, \mathbf{u})$ can be viewed as an accurate eigenpair.

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

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

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

The error can not be computed.

Alternative: compute the **residual** $\mathbf{r} \equiv \mathbf{Au} - \vartheta \mathbf{u}$.

To ease notation, we scale **u** such that $\|\mathbf{u}\|_2 = 1$.

We will learn that,

If the residual norm $\|\mathbf{r}\|_2$ is small,
then $(\vartheta, \mathbf{u})$ can be viewed as an accurate eigenpair.

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

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

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

The error can not be computed.

Alternative: compute the **residual** $\mathbf{r} \equiv \mathbf{A}\mathbf{u} - \vartheta\mathbf{u}$.

Estimates will be given for $|\lambda - \vartheta|$, $\sin\angle(\mathbf{x}, \mathbf{u})$, $\tan\angle(\mathbf{x}, \mathbf{u})$

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

In practice: $\mathbf{A}$ will be perturbed:

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

where the perturbation Δ is $n \times n$ and (hopefully) small.

Perturbations come from

1) rounding errors
2) approximation errors
3) analysis technique

Note. $([\mathbf{A} + \Delta] - \Delta)\mathbf{x} = \lambda\,\mathbf{x}$:
results for $\mathbf{x}$ and $\mathbf{u}$ are interchangeable.

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

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

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

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

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.

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

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

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

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

A perturbation $\boldsymbol{\Delta}$ of $\mathbf{A}$ such that

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

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

in this context the ‘error’ is called the ‘forward error’.

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

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** $\mathbf{r} \equiv \mathbf{A}\mathbf{u} - \vartheta\mathbf{u}$.

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** $\mathbf{r} \equiv \mathbf{A}\mathbf{u} - \vartheta\mathbf{u}$ is small.

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

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

Note. • $\mathbf{r}\mathbf{u}^*$ is of rank 1.

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

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

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

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

$\rho(\mathbf{u}) \equiv \frac{\mathbf{u}^*\mathbf{A}\mathbf{u}}{\mathbf{u}^*\mathbf{u}}$ 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 $\mathbf{\Delta} \equiv \mathbf{r}\mathbf{u}^*$, we have

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

$\mathbf{y} \neq \mathbf{0}$ is a **left** eigenvector if $\mathbf{y}^*\mathbf{A} = \lambda\mathbf{y}^*$.

$\mathbf{x}$ with $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ is a **right** eigenvector.

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

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

$\mathbf{w}$ approximate left eigenvector, $\|\mathbf{w}\|_2 = 1$, $\mathbf{s}^* \equiv \mathbf{w}^*\mathbf{A} - \vartheta\mathbf{w}^*$ and ϑ the associated **two-sided Rayleigh quotient**:

$$\vartheta \equiv \frac{\mathbf{w}^*\mathbf{A}\mathbf{u}}{\mathbf{w}^*\mathbf{u}}.$$

Proposition. With $\Delta \equiv \mathbf{r}\mathbf{u}^* + \mathbf{w}\mathbf{s}^*$, we have that

$$(\mathbf{A} - \Delta)\mathbf{u} = \vartheta\mathbf{u} \quad \& \quad \mathbf{w}^*(\mathbf{A} - \Delta) = \vartheta\mathbf{w}^*$$

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

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

$\mathbf{w}$ approximate left eigenvector, $\|\mathbf{w}\|_2 = 1$, $\mathbf{s}^* \equiv \mathbf{w}^*\mathbf{A} - \vartheta\mathbf{w}^*$ and ϑ the associated **two-sided Rayleigh quotient**:

$$\vartheta \equiv \frac{\mathbf{w}^*\mathbf{A}\mathbf{u}}{\mathbf{w}^*\mathbf{u}}.$$

Proposition. With $\Delta \equiv \mathbf{r}\mathbf{u}^* + \mathbf{w}\mathbf{s}^*$, we have that

$$(\mathbf{A} - \Delta)\mathbf{u} = \vartheta\mathbf{u} \quad \& \quad \mathbf{w}^*(\mathbf{A} - \Delta) = \vartheta\mathbf{w}^*$$

Note. If $\mathbf{A}^* = \mathbf{A}$ and $\mathbf{w} = \mathbf{u}$,

then $\Delta = \mathbf{r}\mathbf{u}^* + \mathbf{w}\mathbf{s}^* = \mathbf{r}\mathbf{u}^* + \mathbf{u}\mathbf{r}^*$ is Hermitian.

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 $\mathbf{\Delta} \equiv \mathbf{r}\mathbf{u}^*$, we have

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

If $\|\mathbf{r}\|_2$ is small, then the question whether $(\vartheta, \mathbf{u})$ is an accurate approximate eigenpair depends on how sensitive eigenpairs of $\mathbf{A}$ are to perturbations.

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

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

- 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.

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

The fact that normal matrices have an orthonormal basis of eigenvectors makes the perturbation theory for these matrices much easier with stronger results (and different proofs) than for general matrices.

The results for Hermitian matrices $\mathbf{A}$ are particularly nice, also due to the fact that the eigenvalues for these matrices have a natural ordering:

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n,$$

where $\lambda_j = \lambda_j(\mathbf{A})$ are the eigenvalues of $\mathbf{A}$ counted according to multiplicity.

These properties are reflected in, for instance, the Courant–Fischer theorem on the next transparency.

We will give results for general matrices, as well as the corresponding results for normal or Hermitian matrices.

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

Theorem [Courant–Fischer] If $\lambda_1 \leq \ldots \leq \lambda_n$, 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 $\mathcal{W}$.

This theorem generalises the fact that $\rho(\mathbf{u})$ is a convex mean of all λ_i and therefore,

$$\lambda_1 = \min_{\mathbf{u}} \rho(\mathbf{u})$$

($\mathbf{u} = \mathbf{x}_1$, leads to the equality).

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

Theorem [Courant–Fischer] If $\lambda_1 \leq \ldots \leq \lambda_n$, 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 $\mathcal{W}$.

Proof. If $\mathcal{W}$ is i-dimensional, then a dimension argument reveals that

$$\mathcal{W} \cap \text{span}(\{\mathbf{x}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n\}) \neq \{\mathbf{0}\}$$

If $\mathbf{w}$ is a non-zero vector in this intersection, then $\rho(\mathbf{w})$ is a convex mean of $\lambda_i, \ldots, \lambda_n$, whence, greater than or equal to λ_i.

With $\mathcal{W} = \text{span}(\mathbf{x}_1, \ldots, \mathbf{x}_i)$ and $\mathbf{w} = \mathbf{x}_i$ we have the equality.

See also Exercise 12.2.

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

Theorem [Courant–Fischer] If $\lambda_1 \leq \ldots \leq \lambda_n$, 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**:

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

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

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

Theorem [Courant–Fischer] If $\lambda_1 \leq \ldots \leq \lambda_n$, 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**:

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

Proof. Apply CF with $\mathcal{W} \subset \mathrm{span}(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1})$. [Exercise 12.4]

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

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.

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

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

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

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

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

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

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

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 $\{\lambda_j \mid j \neq j_0\}$.
In case $\mathbf{x} = \mathbf{x}_1$ we have that $\rho(\mathbf{z}) \geq \lambda_2$ (Courant–Fischer).

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

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

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

- 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.

Smooth perturbations

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

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

We put the "smallness" of the perturbation in τ:

if $\mathbf{A} - \Delta$ is the perturbed matrix with $\|\Delta\|_2 \ll \|\mathbf{A}\|_2$, then we put

$$\epsilon \equiv \|\Delta\|_2, \qquad \mathbf{E} \equiv \frac{\Delta}{\|\Delta\|_2}.$$

Hence, $\quad \mathbf{A}(\epsilon) = \mathbf{A} - \epsilon\mathbf{E} = \mathbf{A} - \Delta$,

$\|\mathbf{E}\|_2 = 1$, and

$\epsilon \ll \|\mathbf{A}\|_2$.

Smooth perturbations

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

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

We put the "smallness" of the perturbation in τ:

if $\mathbf{A} - \Delta$ is the perturbed matrix with $\|\Delta\|_2 \ll \|\mathbf{A}\|_2$, then we put

$$\epsilon \equiv \|\Delta\|_2, \qquad \mathbf{E} \equiv \frac{\Delta}{\|\Delta\|_2}.$$

For a backward error as $\mathbf{ru}^*$:

$$\epsilon\,\mathbf{E} = \Delta = \mathbf{ru}^*, \quad \epsilon = \|\mathbf{r}\|_2, \quad \mathbf{A} - \|\mathbf{r}\|_2\,\mathbf{E}.$$

Smooth perturbations

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

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 $(\tau \in \mathbb{C})$.

Proof. Apply complex function theory to asses multiplicity.

Examples

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

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

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 simple eigenvalue of $\mathbf{A}(\tau)$ at $\tau = 0$, then $\lambda_i(\tau)$ can be expressed as a power series

$$\lambda_i(\tau) = \lambda + \sum_{j=1}^{\infty} \alpha_j \tau^j \qquad (\tau \approx 0).$$

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 $\omega \equiv e^{2\pi i/p}$, $p \leq \text{mult}(\lambda)$.

Smooth perturbations

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

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

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

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

$$\begin{aligned} \tan\angle(\mathbf{x}, \mathbf{u}) &= \|(\widetilde{\mathbf{A}} - \lambda\mathbf{I})^{-1}\,\Delta\,\mathbf{x}\|_2 + \mathcal{O}(\|\Delta\|_2^2) \\ &\lesssim \|(\widetilde{\mathbf{A}} - \lambda\mathbf{I})^{-1}\|_2\,\|\Delta\|_2 \end{aligned}$$

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

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

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

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

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.

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

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

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

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

$\Delta = \mathbf{r}\mathbf{u}^*$. **Corollary.**

$$\tan \angle(\mathbf{x}, \mathbf{u}) \leq \|(\widetilde{\mathbf{A}} - \lambda \mathbf{I})^{-1}\|_2 \, \|\mathbf{r}\|_2 + \mathcal{O}(\|\mathbf{r}\|_2^2)$$

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

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

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

$\Delta = \mathbf{r}\mathbf{u}^*$. **Theorem.** If $\mathbf{A}$ is normal, then

$$\sin\angle(\mathbf{x}, \mathbf{u}) \leq \frac{\|\mathbf{r}\|_2}{\gamma\,\cos\angle(\mathbf{x}, \mathbf{u})}.$$

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

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

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

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

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

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

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

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.

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

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

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

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

$\Delta = \mathbf{r}\mathbf{u}^*$. **Corollary.** For a $\lambda \in \Lambda(\mathbf{A})$, we have that

$$|\vartheta - \lambda| \leq \frac{\|\mathbf{r}\|_2}{\mathbf{y}^*\mathbf{x}} + \mathcal{O}(\|\mathbf{r}\|_2^2).$$

Note that here ϑ is not required to be $\rho(\mathbf{u})$,
whence $\mathbf{y}^*\Delta\mathbf{x} = \mathbf{u}^*\mathbf{r}\mathbf{u}^*\mathbf{u}$ need not be 0.

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

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

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

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

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

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

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

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

$\Delta = \mathbf{r}\mathbf{u}^*$. **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; $\mathbf{y}^*\Delta\mathbf{x} = \mathbf{u}^*\mathbf{r}\mathbf{u}^*\mathbf{u} = 0$).

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

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

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

$\Delta = \mathbf{r}\mathbf{u}^*$. **Theorem [Bauer-Fike].** For general square $\mathbf{A}$

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

assuming there is a basis $\mathbf{X}$ of eigenvectors of $\mathbf{A}$ and

$$\mathcal{C}_E \equiv \|\mathbf{X}\|_2\,\|\mathbf{X}^{-1}\|_2$$

is the conditioning of this basis.

The conditioning of an eigenvalue

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

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

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

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

Theorem [Henrici]. Let $\mathbf{A} = \mathbf{QSQ}^*$ be the Schur decomposition of $\mathbf{A}$: $\mathbf{Q}$ unitary and $\mathbf{S} = \mathbf{D} + \mathbf{S}'$ where $\mathbf{D}$ is diagonal and $\mathbf{S}'$ is strictly upper triangular. Then

$$|\vartheta - \lambda| \leq \max(\epsilon, \epsilon^{\frac{1}{p}}) \quad \text{for some} \quad \lambda \in \Lambda(\mathbf{A})$$

where $\epsilon \equiv \|\mathbf{r}\|_2 \sum_{k=0}^{p-1} \|\mathbf{S}'\|_2^k$ and p such that $(\mathbf{S}')^p = 0$.

$\|\mathbf{S}'\|_2$ is $\mathbf{A}$'s **departure of normality**.
The departure of normality relates to $\|\mathbf{A}^*\mathbf{A} - \mathbf{A}\mathbf{A}^*\|_2$.

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

In this new basis setting,

$$\mathbf{x}(0) = \mathbf{x} = \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{x}(\tau) = \begin{bmatrix} 1 \\ \mathbf{z}_\tau \end{bmatrix} = (1, \mathbf{z}_\tau^\top)^\top.$$

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 $(1, \mathbf{0}^\top)^\top$ and the perturbed eigenvector $(1, \mathbf{z}_\tau^\top)^\top$.

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 $(1, \mathbf{0}^\top)^\top$ and the perturbed eigenvector $(1, \mathbf{z}_\tau^\top)^\top$.

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

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 $(1, \mathbf{0}^\top)^\top$ and the perturbed eigenvector $(1, \mathbf{z}_\tau^\top)^\top$.

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

Hence, for $\tau \to 0$,

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

Proof. Express λ_τ and $\mathbf{z}_\tau$ as a series in powers of τ, plug in and equate τ^j-terms for subsequent $j = 0, 1, 2, \ldots$.

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

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

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

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

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

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 $\lambda(0) = \lambda$ 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 $\mathcal{C}_2(\mathbf{V})$.

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

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

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

Proof. $\mathbf{V} = [\mathbf{x}, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ with $(\mathbf{v}_2, \ldots, \mathbf{v}_n)$ orthonormal basis $\mathbf{x}^\perp$.

Then, $\mathbf{A}\mathbf{V}\mathbf{e}_1 = \mathbf{A}\mathbf{x} = \lambda\mathbf{x} = \lambda\mathbf{V}\mathbf{e}_1$. Hence, $\mathbf{e}_1^*(\mathbf{V}^{-1}\mathbf{A}\mathbf{V})\mathbf{e}_1 = \lambda$.

Apply $(*)$ to $\mathbf{V}^{-1}\mathbf{A}\mathbf{V}$, to see that

$$\tan\angle(\mathbf{x}, \mathbf{x}(\mathbf{A} - \tau\mathbf{E})) = \|(\mathbf{A}_1 - \lambda\mathbf{I})^{-1}(\tau\widetilde{\mathbf{r}})\|_2 + \mathcal{O}(\tau^2).$$

Use, $\|\widetilde{\mathbf{r}}\|_2 \le \|\mathbf{V}^{-1}\mathbf{E}\mathbf{V}\|_2 = \|\mathbf{E}\|_2$. $\mathbf{A}_1$ is the matrix of $\widetilde{\mathbf{A}}$ wrt $\mathbf{v}_2, \ldots, \mathbf{v}_n$.

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

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

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

Proof. $\mathbf{V} = [\mathbf{x}, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ with $(\mathbf{v}_2, \ldots, \mathbf{v}_n)$ orthonormal basis $\mathbf{y}^\perp$.

$\mathbf{A}\mathbf{V}\mathbf{e}_1 = \mathbf{A}\mathbf{x} = \lambda\mathbf{x} = \lambda\mathbf{V}\mathbf{e}_1$ implies that $\mathbf{V}^{-1}\mathbf{A}\mathbf{V}\mathbf{e}_1 = \lambda\mathbf{e}_1$.

Since $\mathbf{e}_1^* = \frac{1}{\mathbf{y}^*\mathbf{x}}\mathbf{y}^*\mathbf{V}$, we have $\mathbf{e}_1^*\mathbf{V}^{-1} = \frac{1}{\mathbf{y}^*\mathbf{x}}\mathbf{y}^*$.

Therefore, $\mathbf{e}_1^*\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \frac{1}{\mathbf{y}^*\mathbf{x}}\mathbf{y}^*\mathbf{A}\mathbf{V} = \frac{\lambda}{\mathbf{y}^*\mathbf{x}}\mathbf{y}^*\mathbf{V} = \lambda\mathbf{e}_1^*$.

Now, apply $(*)$ to $\mathbf{V}^{-1}\mathbf{A}\mathbf{V}$, to see that

$$|\lambda(\mathbf{A} - \tau\mathbf{E}) - \lambda(\mathbf{A})| = |\lambda(\tau) - \lambda| = |\tau\nu| + \mathcal{O}(\tau^2).$$

The results follows from $\quad \nu = \mathbf{e}_1^*(\mathbf{V}^{-1}\tau\mathbf{E}\mathbf{V})\mathbf{e}_1 = \dfrac{1}{\mathbf{y}^*\mathbf{x}}\mathbf{y}^*(\tau\mathbf{E})\mathbf{x}.$

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

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

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

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 $\mathbf{A} - \vartheta\mathbf{I} \leq \epsilon$

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

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 $\mathbf{c} \equiv \mathbf{A}\mathbf{u}$ is exactly $\mathbf{c} = (\mathbf{A} + \Delta)\mathbf{u}$ for some small perturbation Δ.
- If $\mathbf{r} = \mathbf{A}\mathbf{u} - \vartheta \mathbf{u} \;\Rightarrow\; \vartheta \in \Lambda_\epsilon(\mathbf{A})$ for $\epsilon \geq \|\mathbf{r}\|_2$.