Introduction to Arnoldi method

SF2524 - Matrix Computations for Large-scale Systems

KTH Royal Institute of Technology (Elias Jarlebring)

2014-11-07

Agenda lecture 2

- Introduction to Arnoldi method
- Gram-Schmidt - efficiency and roundoff errors
- Derivation of Arnoldi method
- (Convergence characterization)

Main eigenvalue algorithms in this course

- Fundamental eigenvalue techniques (Lecture 1)
- Arnoldi method (Lecture 2-3).

- QR-method (Lecture 8-9).

Main eigenvalue algorithms in this course

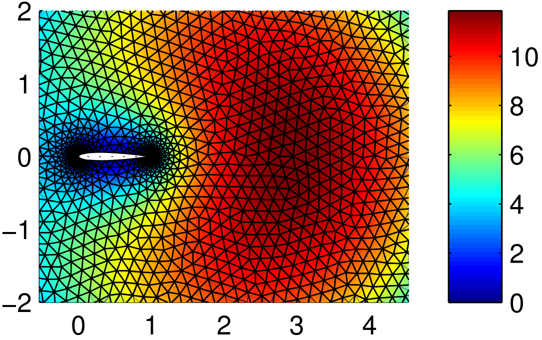
- Fundamental eigenvalue techniques (Lecture 1)
- Arnoldi method (Lecture 2-3).
  Typically suitable when
  - we are interested in a small number of eigenvalues,
  - the matrix is large and sparse
  - Currently solvable size on desktop $m \sim 10^6$ (depending on structure)
- QR-method (Lecture 8-9).
  Typically suitable when
  - we want to compute all eigenvalues,
  - the matrix does not have any particular easy structure.
  - Solvable size on desktop $m \sim 1000$.

Finite-element method example (slide 1/2)

Laplace eigenvalue problem

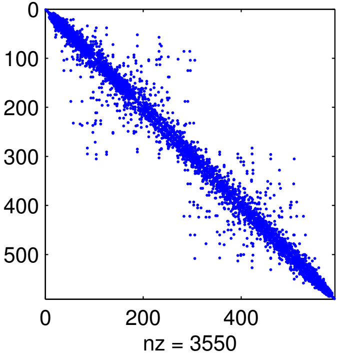
$$\Delta u = \lambda u$$

with Dirichlet boundary conditions outside wing profile.



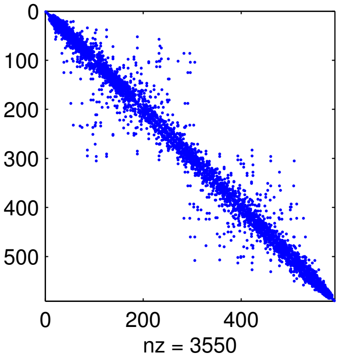
Finite-element method example (slide 2/2)

$\Rightarrow$ Problem: Compute eigenvalues of matrix $A \in \mathbb{R}^{m\times m}$ where m large but few non-zero elements:



Finite-element method example (slide 2/2)

$\Rightarrow$ Problem: Compute eigenvalues of matrix $A \in \mathbb{R}^{m\times m}$ where m large but few non-zero elements:

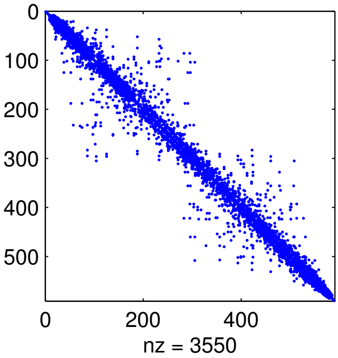


In MATLAB with very efficient QR-method:

```
tic; eig(full(A)); toc
```

Finite-element method example (slide 2/2)

$\Rightarrow$ Problem: Compute eigenvalues of matrix $A \in \mathbb{R}^{m\times m}$ where m large but few non-zero elements:



In MATLAB with very efficient QR-method:

```
tic; eig(full(A)); toc
```

More than 10 hours $\Rightarrow$ we need a different method.

Idea of Arnoldi method (slide 1/3)

Eigenvalue problem

$$Ax = \lambda x.$$

Idea of Arnoldi method (slide 1/3)

Eigenvalue problem

$$Ax = \lambda x.$$

Suppose $Q = (q_1, \ldots, q_n) \in \mathbb{R}^{m \times n}$ and

$$x \in \mathrm{span}(q_1, \ldots, q_n)$$

Idea of Arnoldi method (slide 1/3)

Eigenvalue problem

$$Ax = \lambda x.$$

Suppose $Q = (q_1, \ldots, q_n) \in \mathbb{R}^{m \times n}$ and

$$x \in \mathrm{span}(q_1, \ldots, q_n)$$

$\Rightarrow$ There exists $z \in \mathbb{R}^n$ such that

$$AQz = \lambda Qz$$

Idea of Arnoldi method (slide 1/3)

Eigenvalue problem

$$Ax = \lambda x.$$

Suppose $Q = (q_1, \ldots, q_n) \in \mathbb{R}^{m\times n}$ and

$$x \in \mathrm{span}(q_1, \ldots, q_n)$$

$\Rightarrow$ There exists $z \in \mathbb{R}^n$ such that

$$AQz = \lambda Qz$$

and

$$Q^T AQz = \lambda Q^T Qz$$

Idea of Arnoldi method (slide 1/3)

Eigenvalue problem

$$Ax = \lambda x.$$

Suppose $Q = (q_1, \ldots, q_n) \in \mathbb{R}^{m \times n}$ and

$$x \in \mathrm{span}(q_1, \ldots, q_n)$$

$\Rightarrow$ There exists $z \in \mathbb{R}^n$ such that

$$AQz = \lambda Qz$$

and Q orthogonal

$$Q^T AQz = \lambda Q^T Qz = \lambda z$$

Idea of Arnoldi method (slide 2/3)

Rayleigh-Ritz procedure

Idea of Arnoldi method (slide 2/3)

Rayleigh-Ritz procedure

Let $Q \in \mathbb{R}^{m \times n}$ be orthogonal basis of some subspace.

Idea of Arnoldi method (slide 2/3)

Rayleigh-Ritz procedure

Let $Q \in \mathbb{R}^{m \times n}$ be orthogonal basis of some subspace. Solutions (μ, z) to the eigenvalue problem

$$Q^T AQz = \mu z$$

are called *Ritz pairs*.

Idea of Arnoldi method (slide 2/3)

Rayleigh-Ritz procedure

Let $Q \in \mathbb{R}^{m \times n}$ be orthogonal basis of some subspace. Solutions (μ, z) to the eigenvalue problem

$$Q^T AQz = \mu z$$

are called *Ritz pairs*. The procedure (Rayleigh-Ritz)

- returns an exact eigenvalue if $x \in \operatorname{span} q_1, \ldots, q_n$,

Idea of Arnoldi method (slide 2/3)

Rayleigh-Ritz procedure

Let $Q \in \mathbb{R}^{m \times n}$ be orthogonal basis of some subspace. Solutions (μ, z) to the eigenvalue problem

$$Q^T A Q z = \mu z$$

are called *Ritz pairs*. The procedure (Rayleigh-Ritz)

- returns an exact eigenvalue if $x \in \operatorname{span} q_1, \ldots, q_n$, and
- approximates eigenvalues of A "well" if

$$x \approx \tilde{x} \in \operatorname{span}(q_1, \ldots, q_n).$$

Idea of Arnoldi method (slide 2/3)

Rayleigh-Ritz procedure

Let $Q \in \mathbb{R}^{m\times n}$ be orthogonal basis of some subspace. Solutions (μ, z) to the eigenvalue problem

$$Q^T AQz = \mu z$$

are called *Ritz pairs*. The procedure (Rayleigh-Ritz)

- returns an exact eigenvalue if $x \in \operatorname{span} q_1, \ldots, q_n$, and
- approximates eigenvalues of A "well" if

$$x \approx \tilde{x} \in \operatorname{span}(q_1, \ldots, q_n).$$

What is a good subspace $\operatorname{span}(Q)$?

Idea of Arnoldi method (slide 2/3)

Rayleigh-Ritz procedure

Let $Q \in \mathbb{R}^{m \times n}$ be orthogonal basis of some subspace. Solutions (μ, z) to the eigenvalue problem

$$Q^T AQz = \mu z$$

are called *Ritz pairs*. The procedure (Rayleigh-Ritz)

- returns an exact eigenvalue if $x \in \operatorname{span} q_1, \ldots, q_n$, and
- approximates eigenvalues of A "well" if

$$x \approx \tilde{x} \in \operatorname{span}(q_1, \ldots, q_n).$$

What is a good subspace $\operatorname{span}(Q)$?

Idea of Arnoldi method (slide 3/3)

Recall: Power method approximates the largest eigenvalue well.

Idea of Arnoldi method (slide 3/3)

Recall: Power method approximates the largest eigenvalue well.

Definition: Krylov matrix / subspace

$$K_n(A, q_1) := (q_1, Aq_1, \dots A^{n-1}q_1)$$

$$\mathcal{K}_n(A, q_1) := \operatorname{span}(q_1, Aq_1, \dots A^{n-1}q_1).$$

Idea of Arnoldi method (slide 3/3)

Recall: Power method approximates the largest eigenvalue well.

Definition: Krylov matrix / subspace

$$K_n(A, q_1) := (q_1, Aq_1, \ldots A^{n-1}q_1)$$

$$\mathcal{K}_n(A, q_1) := \mathrm{span}(q_1, Aq_1, \ldots A^{n-1}q_1).$$

Justification of Arnoldi method

- Use Rayleigh-Ritz on $Q = (q_1, \ldots, q_n)$ where

$$\mathrm{span}(q_1, \ldots, q_n) = \mathcal{K}_n(A, q_1)$$

Idea of Arnoldi method (slide 3/3)

Recall: Power method approximates the largest eigenvalue well.

Definition: Krylov matrix / subspace

$$K_n(A, q_1) := (q_1, Aq_1, \ldots A^{n-1}q_1)$$

$$\mathcal{K}_n(A, q_1) := \mathrm{span}(q_1, Aq_1, \ldots A^{n-1}q_1).$$

Justification of Arnoldi method

- Use Rayleigh-Ritz on $Q = (q_1, \ldots, q_n)$ where

$$\mathrm{span}(q_1, \ldots, q_n) = \mathcal{K}_n(A, q_1)$$

- Arnoldi method is a "clever" procedure to construct $H_n = V^T AV$.

Idea of Arnoldi method (slide 3/3)

Recall: Power method approximates the largest eigenvalue well.

Definition: Krylov matrix / subspace

$$K_n(A, q_1) := (q_1, Aq_1, \dots A^{n-1} q_1)$$

$$\mathcal{K}_n(A, q_1) := \operatorname{span}(q_1, Aq_1, \dots A^{n-1} q_1).$$

Justification of Arnoldi method

- Use Rayleigh-Ritz on $Q = (q_1, \dots, q_n)$ where

$$\operatorname{span}(q_1, \dots, q_n) = \mathcal{K}_n(A, q_1)$$

- Arnoldi method is a "clever" procedure to construct $H_n = V^T AV$.
- "Clever":
  (1) We exploit sparsity (we will only need matrix-vector product)

Idea of Arnoldi method (slide 3/3)

Recall: Power method approximates the largest eigenvalue well.

Definition: Krylov matrix / subspace

$$K_n(A, q_1) := (q_1, Aq_1, \dots A^{n-1} q_1)$$

$$\mathcal{K}_n(A, q_1) := \mathrm{span}(q_1, Aq_1, \dots A^{n-1} q_1).$$

Justification of Arnoldi method

- Use Rayleigh-Ritz on $Q = (q_1, \dots, q_n)$ where

$$\mathrm{span}(q_1, \dots, q_n) = \mathcal{K}_n(A, q_1)$$

- Arnoldi method is a “clever” procedure to construct $H_n = V^T AV$.
- “Clever”:
  (1) We exploit sparsity (we will only need matrix-vector product)
  (2) We expand V with one row in each iteration
  $\Rightarrow$ Iterate until we are happy.

Idea of Arnoldi method (slide 3/3)

Recall: Power method approximates the largest eigenvalue well.

Definition: Krylov matrix / subspace

$$K_n(A, q_1) := (q_1, Aq_1, \ldots A^{n-1}q_1)$$

$$\mathcal{K}_n(A, q_1) := \mathrm{span}(q_1, Aq_1, \ldots A^{n-1}q_1).$$

Justification of Arnoldi method

- Use Rayleigh-Ritz on $Q = (q_1, \ldots, q_n)$ where

$$\mathrm{span}(q_1, \ldots, q_n) = \mathcal{K}_n(A, q_1)$$

- Arnoldi method is a "clever" procedure to construct $H_n = V^T AV$.
- "Clever":
  (1) We exploit sparsity (we will only need matrix-vector product)
  (2) We expand V with one row in each iteration
  $\Rightarrow$ Iterate until we are happy.

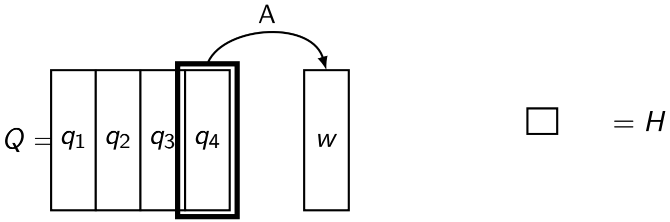
Arnoldi method graphically

Graphical illustration of algorithm:

$$Q = \begin{array}{|c|c|c|c|} \hline q_1 & q_2 & q_3 & q_4 \\ \hline \end{array} \qquad\qquad \square \quad = H$$

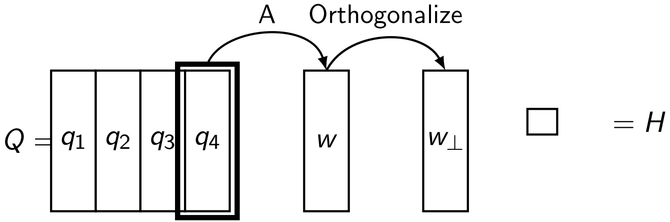
Arnoldi method graphically

Graphical illustration of algorithm:



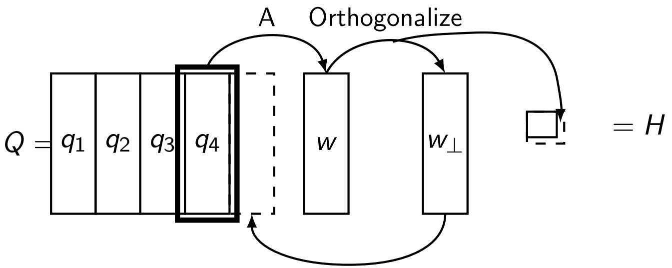
Arnoldi method graphically

Graphical illustration of algorithm:



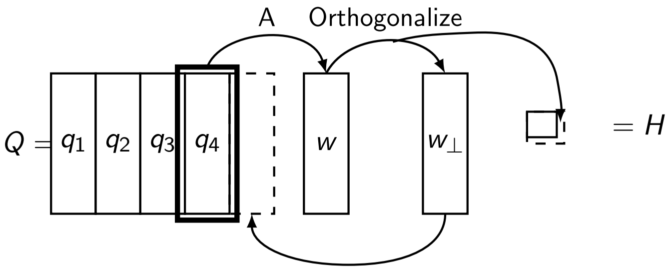
Arnoldi method graphically

Graphical illustration of algorithm:



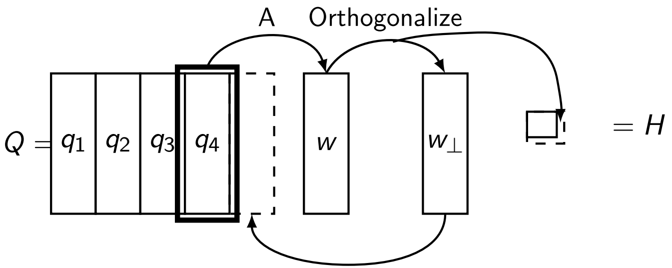
Arnoldi method graphically

Graphical illustration of algorithm:



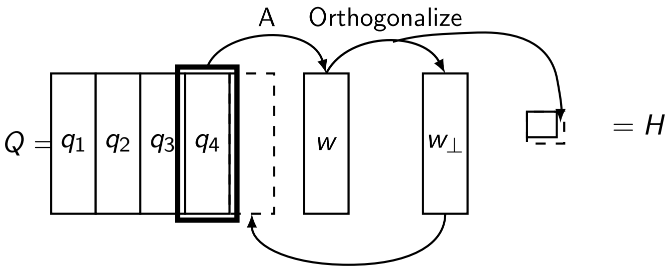
Arnoldi method graphically

Graphical illustration of algorithm:



Arnoldi method graphically

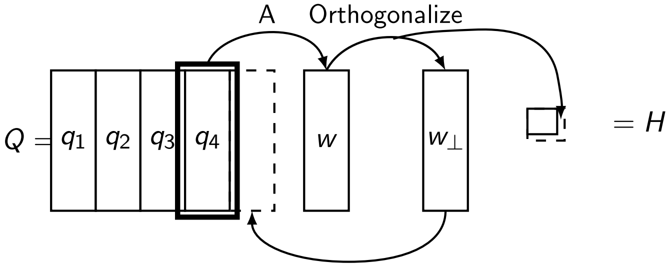
Graphical illustration of algorithm:



After iteration: Take eigenvalues of H as approximate eigenvalues.

Arnoldi method graphically

Graphical illustration of algorithm:

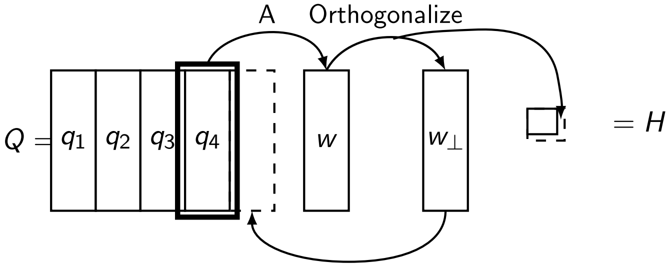


After iteration: Take eigenvalues of H as approximate eigenvalues.

* show arnoldi.m and Hessenberg matrix in matlab *

Arnoldi method graphically

Graphical illustration of algorithm:



After iteration: Take eigenvalues of H as approximate eigenvalues.

* show arnoldi.m and Hessenberg matrix in matlab *

We will now...

(1) derive a good orthogonalization procedure: variants of Gram-Schmidt,

(2) show that Arnoldi generates a Rayleigh-Ritz approximation,

(3) characterize the convergence.