Lecture 33

The Arnoldi Iteration

NLA Reading Group Spring ’13

by Hakan Güldaş

	$A = QR$	$A = QHQ^*$
orthogonal structuring	Householder	Householder
structured orthogonalization	Gram–Schmidt	Arnoldi

Arnoldi is analogue of Gram-Scmidt for similarity transformations to Hessenberg form rather than QR factorization.

It can be stopped part-way, partial reduction to Hessenberg form.

Mechanics of the Arnoldi Iteration

We want to transform *A* unitarily into upper Hessenberg form :

$$A = QHQ^*, \text{ or } AQ = QH.$$

$$\left[\begin{array}{c} \\ \\ A \\ \\ \\ \end{array}\right]\left[\begin{array}{c|c|c} & & \\ q_1 & \cdots & q_n \\ & & \end{array}\right] = \left[\begin{array}{c|c|c} & & \\ q_1 & \cdots & q_{n+1} \\ & & \end{array}\right]\left[\begin{array}{ccc} h_{11} & \cdots & h_{1n} \\ h_{21} & & \vdots \\ & \ddots & \\ & & h_{n+1,n} \end{array}\right]$$

$$AQ_n = Q_{n+1}\tilde{H}_n,$$

$$Aq_n = h_{1n}q_1 + \cdots + h_{nn}q_n + h_{n+1,n}q_{n+1} \quad (33.4)$$

This is a recurrence relation for q_n

Algorithm 33.1. Arnoldi Iteration

$b = \text{arbitrary}, \quad q_1 = b/\|b\|$

for $n = 1, 2, 3, \ldots$

$\quad v = Aq_n$

$\quad$ **for** $j = 1$ **to** n

$\quad\quad h_{jn} = q_j^* v$

$\quad\quad v = v - h_{jn} q_j$

$\quad h_{n+1,n} = \|v\|$ [see Exercise 33.2 concerning $h_{n+1,n} = 0$]

$\quad q_{n+1} = v/h_{n+1,n}$

This is modified Gram-Schmidt that implements 33.4

QR Factorization of a Krylov Matrix

Krylov subspaces generated by *A* and *b*

$$\mathcal{K}_n = \langle b, Ab, \ldots, A^{n-1}b \rangle = \langle q_1, q_2, \ldots, q_n \rangle \subseteq \mathbb{C}^m$$

n × *m Krylov matrix* :

$$K_n = \left[\begin{array}{c|c|c|c} b & Ab & \cdots & A^{n-1}b \end{array} \right]$$

$$K_n = Q_n R_n$$

Arnoldi iteration is based upon the QR factorization of the matrix with columns $b, Ab, \ldots, A^{n-1}b$

Simultaneous iteration and the QR algorithm is based upon the QR factorization of the matrix with columns $A^n e_1, A^n e_2, \ldots A^n e_m$

	quasi-direct	iterative
straightforward but unstable	simultaneous iteration	(33.6)–(33.7)
subtle but stable	QR algorithm	Arnoldi

Projection onto Krylov Subspaces

$$Q_n^* Q_{n+1} \text{ is the } n \times (n+1) \text{ identity}$$

$$Q_n^* Q_{n+1} \tilde{H}_n \text{ is the } n \times n \text{ Hessenberg matrix}$$

$$\text{obtained by removing the last row of } \tilde{H}_n$$

$$H_n = \begin{bmatrix} h_{11} & & \cdots & h_{1n} \\ h_{21} & h_{22} & & \\ & \ddots & \ddots & \vdots \\ & & h_{n,n-1} & h_{nn} \end{bmatrix}$$

$$H_n = Q_n^* \, A \, Q_n$$

$\{q_1, \ldots, q_n\}$ is a basis for $\mathcal{K}_n$

$Q_n^* A\, Q_n$ is the representation of *A* in the basis $\{q_1, \ldots, q_n\}$

Theorem 33.1. *The matrices Q_n generated by the Arnoldi iteration are reduced QR factors of the Krylov matrix* (33.6)*:*

$$K_n = Q_n R_n. \tag{33.11}$$

The Hessenberg matrices H_n are the corresponding projections

$$H_n = Q_n^* A\, Q_n, \tag{33.12}$$

and the successive iterates are related by the formula

$$A Q_n = Q_{n+1} \tilde{H}_n. \tag{33.13}$$

Lecture 34

How Arnoldi Locates Eigenvalues

NLA Reading Group Spring ’13

by Hakan Güldaş

Computing Eigenvalues by the Arnoldi Algorithm

- Compute H_n by Arnoldi algorithm
- Compute eigenvalues of H_n by standard methods (such as QR)
- These eigenvalues converge to extreme eigenvalues of A

Arnoldi and Polynomial Approximation

Let x be a vector in the Krylov subspace $\mathcal{K}_n$ (33.5). Such an x can be written as a linear combination of powers of A times b:

$$x = c_0 b + c_1 A b + c_2 A^2 b + \cdots + c_{n-1} A^{n-1} b. \qquad (34.1)$$

$$P^n = \{\text{monic polynomials of degree } n\}.$$

Arnoldi/Lanczos Approximation Problem. Find $p^n \in P^n$ such that

$$\| p^n(A) b \| = \text{minimum}. \qquad (34.3)$$

Theorem 34.1. *As long as the Arnoldi iteration does not break down (i.e., K_n is of full rank n), (34.3) has a unique solution p^n, namely, the characteristic polynomial of H_n.*

Invariance Properties

Theorem 34.2. *Let the Arnoldi iteration be applied to a matrix* $A \in \mathbb{C}^{m \times m}$ *as described above.*

Translation-invariance. *If* A *is changed to* $A + \sigma I$ *for some* $\sigma \in \mathbb{C}$, *and* b *is left unchanged, the Ritz values* $\{\theta_j\}$ *at each step change to* $\{\theta_j + \sigma\}$.

Scale-invariance. *If* A *is changed to* σA *for some* $\sigma \in \mathbb{C}$, *and* b *is left unchanged, the Ritz values* $\{\theta_j\}$ *change to* $\{\sigma\theta_j\}$.

Invariance under unitary similarity transformations. *If* A *is changed to* UAU^* *for some unitary matrix* U, *and* b *is changed to* Ub, *the Ritz values* $\{\theta_j\}$ *do not change.*

In all three cases the Ritz vectors, *namely the vectors* $Q_n y_j$ *corresponding to the eigenvectors* y_j *of* H_n, *do not change under the indicated transformation.*

How Arnoldi Locates Eigenvalues

If one's aim is to find a polynomial p^n with the property that $p^n(A)$ is small, an effective means to that end may be to pick p^n to have zeros close to the eigenvalues of A.

Consider an extreme case. Suppose that A is diagonalizable and has only $n \ll m$ distinct eigenvalues, hence a minimal polynomial of degree n. Then from Theorem 34.1 it is clear that after n steps, all of these eigenvalues will be found exactly, at least if the vector b contains components in directions associated with every eigenvalue. Thus after n steps, the Arnoldi iteration has computed the minimal polynomial of A exactly.

Arnoldi Lemniscates

A *lemniscate* is a curve or collection of curves

$$\{z \in \mathbb{C} : \ |p(z)| = C\},$$

If we replace p above by p^n and put

$$C = \frac{||p^n(A)b||}{||b||},$$

this lemniscate is called Arnoldi lemniscate.
As the iteration number n increases, components of these lemniscates typically appear which surround the extreme eigenvalues of A and then shrink rapidly to a point, namely eigenvalues itself.

Geometric Convergence

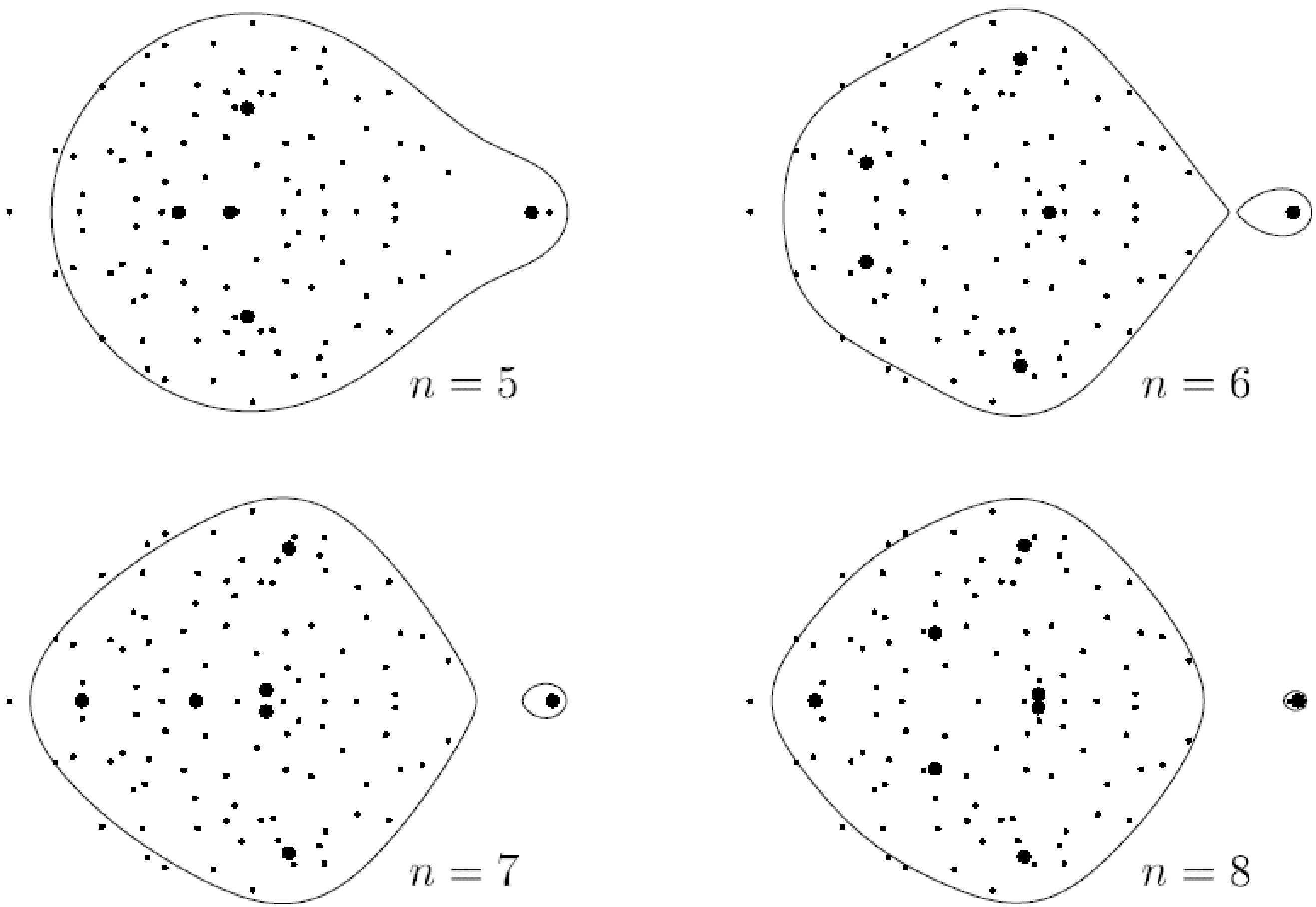

Figure 34.3. *Arnoldi lemniscates (34.4)–(34.5) at steps $n = 5, 6, 7, 8$ for the same matrix A. The small dots are the eigenvalues of A, and the large dots are the eigenvalues of H_n, i.e., the Ritz values. One component of the Arnoldi lemniscate first "swallows" the outlier eigenvalue, and in subsequent iterations it then shrinks to a point at a geometric rate.*

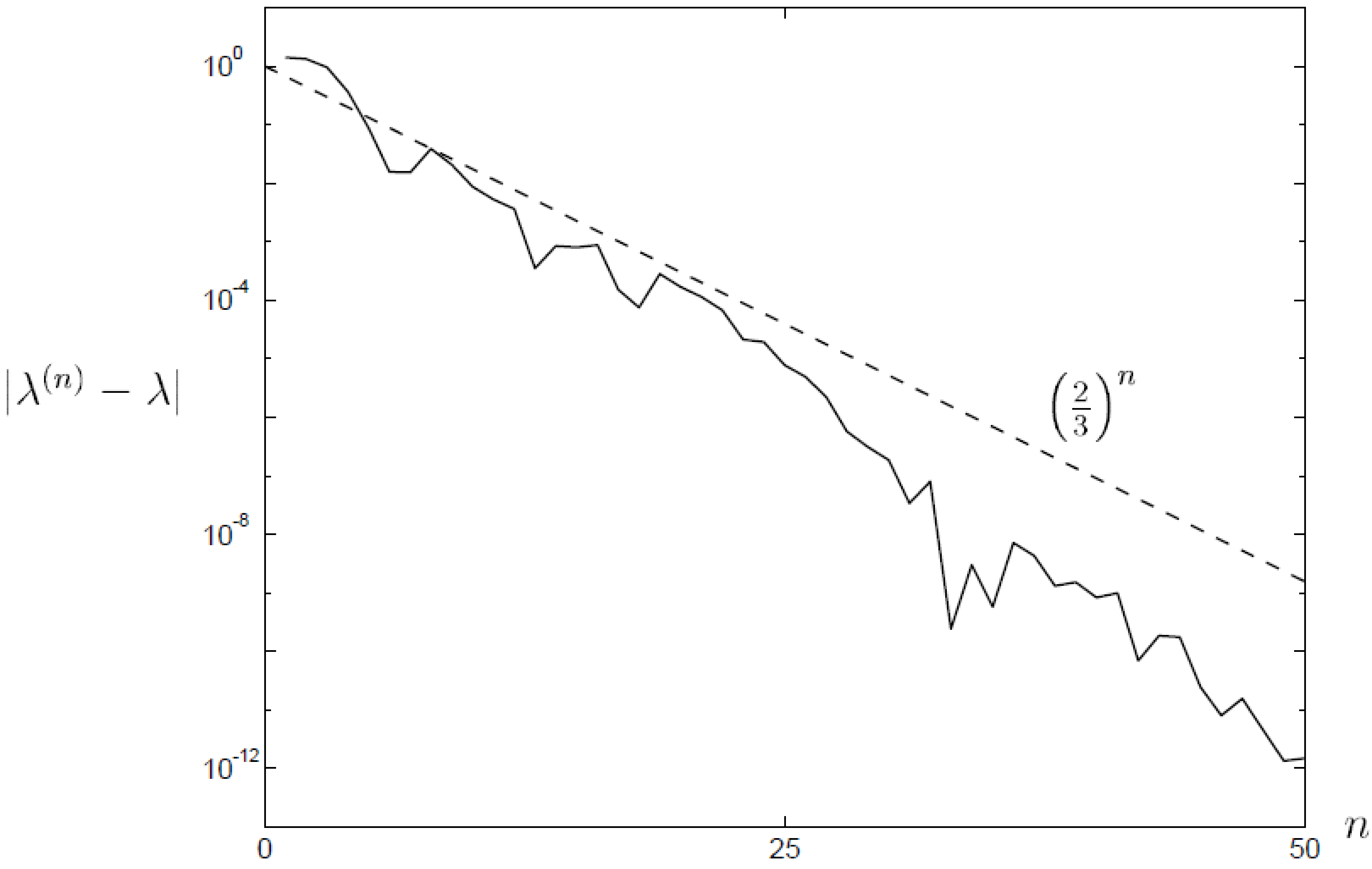

Figure 34.4. *Convergence of the rightmost Arnoldi eigenvalue estimate.*

Figure 34.3 approximates the rate

$$|\lambda^{(n)} - \lambda| \approx \left(\frac{2}{3}\right)^n .$$

Consider the polynomial $p(z) = z^{n-1}(z - \tilde{\lambda})$, where $\tilde{\lambda}$ is some number close to λ.

At each of the eigenvalues of A in the unit disk, $|p(z)|$ is of order 1 or smaller. At $z = \lambda$, however, it has magnitude

$$|p(\lambda)| \approx \left(\frac{3}{2}\right)^n |\tilde{\lambda} - \lambda|$$

(this would be an equality if λ were exactly equal to 3/2). When n is large, $(3/2)^n$ is huge. For this number also to be of order 1, $|\tilde{\lambda} - \lambda|$ must be small enough to balance it, that is, of order $(2/3)^n$, as in (34.6).