Mathematics, Pusan National University

# Numerical Linear Algebra

# Lecture 33. The Arnoldi Iteration

Taehyeong Kim
th_kim@pusan.ac.kr

January 4, 2021

# Content

# Introduction



In this lecture, ours will be to consider the Arnoldi process, a Gram-Schmidt-style iteration for transforming a matrix to Hessenberg form.

# The Arnoldi/Gram-Schmidt Analogy



We can summarize the four algorithms just mentioned in a table:

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

## Krylov subspace

In linear algebra, the order-$r$ Krylov subspace generated by an $n$-by-$n$ matrix $A$ and a vector $\mathbf{b}$ of dimension $n$ is the linear subspace spanned by the images of $\mathbf{b}$ under the first $r$ powers of $A$ (starting from $A^0 = I$), that is,

$$\mathcal{K}_r(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}, A^2\mathbf{b}, \ldots, A^{r-1}\mathbf{b}\}$$

# The Arnoldi/Gram-Schmidt Analogy



## Arnoldi iteration

The Arnoldi iteration was invented by W. E. Arnoldi in 1951.[1] In numerical linear algebra, the Arnoldi iteration is an eigenvalue algorithm and an important example of an iterative method. Arnoldi finds an approximation to the eigenvalues and eigenvectors of general (possibly non-Hermitian) matrices by constructing an orthonormal basis of the Krylov subspace, which makes it particularly useful when dealing with large sparse matrices.

The Arnoldi process needs this vector in order to get started. For applications to eigenvalue problems, we typically assume that $\mathbf{b}$ is random. For applications to systems of equations, as considered in later lectures, it will be the right-hand side, or more generally, the initial residual.

# Mechanics of the Arnoldi Iteration



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

Let $Q_n$ be the matrix $m \times n$ matrix whose columns are the first $n$ columns of $Q$:

$$Q_n = \left[ \begin{array}{c|c|c|c} q_1 & q_2 & \cdots & q_n \end{array} \right] \tag{1}$$

And let $\hat{H}_n$ be the $(n+1) \times n$ upper-left section of $H$, which is also a Hessenberg matris:

$$\hat{H}_n = \begin{bmatrix} h_{11} & & \cdots & & h_{1n} \\ h_{21} & h_{22} & & & h_{2n} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & h_{n,n-1} & h_{nn} \\ & & & & h_{n+1,n} \end{bmatrix} \tag{2}$$

# Mechanics of the Arnoldi Iteration



Then we have

$$AQ_n = Q_{n+1}\hat{H}_n \tag{3}$$

that is,

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

The $n$th column of this equation canbe written as follows:

$$Aq_n = h_{1n}q_n + \cdots + h_{nn}q_n + h_{n+1,n}q_{n+1} \tag{4}$$

# Mechanics of the Arnoldi Iteration



The Arnoldi iteration is simply the modified Gram-Schmidt iteration that implement (4). The following algorithm should be compared with Algorithm 8.1

**Algorithm 8.1. Modified Gram–Schmidt**

**for** $i = 1$ **to** $n$

$\quad v_i = a_i$

**for** $i = 1$ **to** $n$

$\quad r_{ii} = \|v_i\|$

$\quad q_i = v_i / r_{ii}$

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

$\quad\quad r_{ij} = q_i^* v_j$

$\quad\quad v_j = v_j - r_{ij} q_i$

# Mechanics of the Arnoldi Iteration

The Arnoldi iteration is simply the modified Gram-Schmidt iteration that implement (4). The following algorithm should be compared with Algorithm 8.1

**Algorithm 33.1. Arnoldi Iteration**

$b =$ arbitrary, $q_1 = b/\|b\|$

**for** $n = 1, 2, 3, \dots$

  $v = Aq_n$

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

    $h_{jn} = q_j^* v$

    $v = v - h_{jn} q_j$

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

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

# QR Factorization of a Krylov Matrix



## Krylov matrix

Krylov subspace is defined as follows:

$$\mathcal{K}_n = \langle b, Ab, \ldots, A^{n-1}b \rangle = \langle q_1, q_2, \ldots, q_n \rangle \subseteq \mathbb{C}^m \tag{5}$$

And let define $K_n$ to be the $m \times n$ Krylov matrix

$$K_n = \left[ \begin{array}{c|c|c|c} b & Ab & \cdots & A^{n-1}b \end{array} \right] \tag{6}$$

Then $K_n$ must have a reduced QR factorization

$$K_n = Q_n R_n \text{ where } Q_n \text{ is the same matrix as above.} \tag{7}$$

# QR Factorization of a Krylov Matrix



| | quasi-direct | iterative |
|---|---|---|
| straightforward but unstable | simultaneous iteration | (6)-(7) |
| subtle but stable | QR algorithm | Arnoldi |

# Projection onto Krylov Subspaces



Another way to view the Arnoldi process is as a computation of projection onto successive Krylov Subspaces.

$$Q_n^* Q_{n+1} : n \times (n+1) \text{ identity}$$
$$\Rightarrow Q_n^* Q_{n+1} \bar{H}_n : n \times n \text{ Hessenberg matrix}$$

$$H_n = \begin{bmatrix} h_{11} & & \cdots & & h_{1n} \\ h_{21} & h_{22} & & & \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \\ & & & h_{n,n-1} & h_{nn} \end{bmatrix} \tag{8}$$

And from (3),

$$H_n = Q_n^* A Q_n \tag{9}$$

# Projection onto Krylov Subspaces



## Rayleigh-Ritz method

The Rayleigh-Ritz method allows for the computation of Ritz pairs $(\tilde{\lambda}_i, \bar{x}_i)$ which approximate the solution to the eigenvalue problem

$$Ax = \lambda x \text{ where } A \in \mathbb{C}^{n\times n}$$

The procedure is as follows:

1. Compute an orthonormal basis $V \in \mathbb{C}^{n\times m}$ approximating the eigenspace corresponding to $m$ eigenvectors.
2. Compute $R = V^* AV$
3. Compute the eigenvalues of $R$ for solving $Rv_i = \tilde{\lambda}_i v_i$
4. Form the Ritz pairs $(\tilde{\lambda}_i, \bar{x}_i) = (\tilde{\lambda}_i, V v_i)$

If a Krylov subspace is used and $A$ is general matrix, then this is the Arnoldi algorithm.

**Theorem 5.1**

*The matrices $Q_n$ generated by the Arnoldi iteration are reduced QR factors of the Krylov matrix (6):*

$$K_n = Q_n R_n \tag{10}$$

*The Hessenberg matrices $H_n$ are the corresponding projections*

$$H_n = Q_n^* A Q_n \tag{11}$$

*and the successive iterates are related by formula*

$$A Q_n = Q_{n+1} \tilde{H}_n \tag{12}$$

# Reference



[1] Walter Edwin Arnoldi. "The principle of minimized iterations in the solution of the matrix eigenvalue problem". In: *Quarterly of applied mathematics* 9.1 (1951), pp. 17–29.

[2] Nicholas J Higham. *Functions of matrices: theory and computation*. SIAM, 2008.

[3] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[4] Gilbert Strang. *Linear algebra and learning from data*. Wellesley-Cambridge Press, 2019.

[5] Lloyd N Trefethen and David Bau III. *Numerical linear algebra*. Vol. 50. Siam, 1997.

Thank you!