# Krylov subspace approximation of rational matrix functions

*Winfried Auzinger*

Institute for Analysis and Scientific Computing
Vienna University of Technology, Austria

www.math.tuwien.ac.at/∼winfried
mailto:w.auzinger@tuwien.ac.at

Innsbruck, May 13–16, 2012
Innovative Time Integration

## Overview

Review of Arnoldi [one-sided Lanczos] process; notation

Krylov approximation of matrix functions

Review: Approximation of matrix exponential

Approximation of rational matrix function

Representation of approximation error

A posteriori error estimation

## Review of Arnoldi [one-sided Lanczos] process; notation

For given matrix $A \in \mathbb{R}^{n\times n}$ [Lanczos: $A = A^T$]

- Start from $v \in \mathbb{R}^n$; build Krylov space $\mathcal{K}_m = \mathcal{K}_m(A, v)$
- Orthonormal Arnoldi vectors $v_i$ span $\mathcal{K}_m$

$$V_{m+1} = \left(\begin{array}{c|c|c||c} \vdots & & \vdots & \vdots \\ v_1 & \cdots & v_m & \bar{v} \\ \vdots & & \vdots & \vdots \end{array}\right) = \left( V_m \,\|\, \bar{v} \right) \in \mathbb{R}^{n\times(m+1)}$$

- $\bar{v} := v_{m+1} \perp \mathcal{K}_m$
- $h_{i,j} = v_i^T A v_j \quad \rightsquigarrow \quad$ upper Hessenberg matrix ($\gamma := h_{m+1,m}$)

$$\bar{H}_m = \left(\begin{array}{ccccc} h_{1,1} & h_{1,2} & h_{1,3} & \dots & h_{1,m} \\ h_{2,1} & h_{2,2} & h_{2,3} & \dots & h_{2,m} \\ & h_{3,2} & h_{3,3} & \dots & h_{3,m} \\ & & \ddots & \ddots & \vdots \\ & & & h_{m,m-1} & h_{m,m} \\ \hline\hline & & & & \gamma \end{array}\right) = \left(\begin{array}{c} H_m \\ \hline\hline \gamma\, e_m^T \end{array}\right) \in \mathbb{R}^{(m+1)\times m}$$

- $H_m \in \mathbb{R}^{m\times m}$: quadratic upper Hessenberg [or symm. tridiagonal]

- Arnoldi $\rightsquigarrow$

$$A\,V_m = V_m\,H_m + \gamma\,\bar{v}\,\mathbf{e}_m^T, \quad H_m = V_m^T A\,V_m, \qquad V_m^T V_m = I_m$$

- FOM approximation ( $\beta = \|v\|_2$ ):

$$u_m = V_m\,\phi(H_m) V_m^T v = \beta\,V_m\,\phi(H_m)\,\mathbf{e}_1 \approx \phi(A)\,v = u_*$$

- Here:
  - Main interest in rational $\phi(A)\,v$
  - This talk: Representation of error $u_m - u_*$
    (first step towards a posteriori error estimation)
  - Not yet computationally realized
  - Application: Rational integrators

## Review: Approximation of matrix exponential

[ Saad; Hochbruck, Lubich, Selhofer; Lopez, Simoncini; Botchev; ... ]

- FOM approximation:

$$u_m = V_m \exp(tH_m) V_m^T v = \beta V_m \exp(tH_m) e_1 \approx \exp(tA)\, v = u_*$$

- Approximate $\exp(tH_m)$, e.g., by Padé
- ¿ *What is the 'residual' of* $u_m$ ?
- ¿ *Representation of error* $e_m = u_m - u_*$ ?
- Consider $u_m = u_m(t) = V_m \xi_m(t)$ as continuous object $\rightsquigarrow$

$$\xi_m'(t) = H_m \xi_m(t), \quad \xi_m(0) = \beta \mathbf{e}_1$$

  $\Rightarrow$ $\xi_m \in \mathbb{R}^m$ is solution of projected IVP

- *Defect* of $u_m(t)$ w.r.t. $u' = A u$ :

$$r_m(t) := u_m'(t) - A u_m(t) = V_m H_m \xi_m(t) - A V_m \xi_m(t)$$
$$= -\gamma \bar{v} \mathbf{e}_m^T \xi_m(t) = -\gamma (\xi_m(t))_m \bar{v} \perp \mathcal{K}_m$$

- For error estimation: Approximate equation for $e_m = u_m - u_*$,

$$e_m'(t) = A e_m(t) + r_m(t)$$

  again by Krylov techniques

  (Details: Van den Eshof, Hochbruck (2006); Botchev (2011))

## Approximation of rational matrix function

Analogous procedure for $\phi(A) = R(A)$, $R(z) = Q(z)^{-1} P(z)$

- $u_* = R(A)\,v \quad \leftrightsquigarrow \quad$ linear system for $u_*$ :
$$Q(A)\,u = P(A)\,v \qquad (*)$$
- FOM approximation for linear system $(*)$ :
$$u_m = V_m\,Q(H_m)^{-1} V_m^T\,P(A)\,v$$
Assume $m >$ degree$(P)$ $\quad \Rightarrow$
$$\underline{u_m} = V_m\,Q(H_m)^{-1} \underbrace{V_m^T V_m}\,P(H_m) V_m^T v$$
$$= V_m\,Q(H_m)^{-1} \overbrace{I_m}\,P(H_m) V_m^T v = V_m\,R(H_m)\,\underbrace{V_m^T v}_{=\beta\,\mathbf{e_1}}$$
$$= \underline{\text{direct FOM approximation for } \phi(A) = R(A)}$$
- $Q(A)$ usually very ill-conditioned
- $\Rightarrow$ Residual w.r.t. $(*)$,
$$r_m = Q(A)\,u_m - P(A)\,v = Q(A)\,u_m - V_m\,P(H_m) V_m^T v$$
is not a reasonable quality measure for $u_m \approx u_*$.
- Remark: In time integrators, typically $R(tA)$, $t =$ step length

# Representation of approximation error

## Review of FOM residual and error for $A u = v$

- Simplest case: $u_* = A^{-1} v \;\; (P(z) = 1,\; Q(z) = z)$

$\rightsquigarrow$

$$\begin{aligned} u_m &= V_m H_m^{-1} V_m^T v = V_m H_m^{-1} \beta \mathbf{e}_1 \\ &= V_m \xi_m, \quad \text{with} \quad H_m \xi_m = \beta \mathbf{e}_1, \quad \xi_m \in \mathbb{R}^m \end{aligned}$$

$\rightsquigarrow$

$$\begin{aligned} r_m = A u_m - v = A V_m \xi_m - v \\ &= (V_m H_m + \gamma \bar{v} \mathbf{e}_m^T) \xi_m - v \\ &= \not{v} + \gamma (\mathbf{e}_m^T \xi_m) \bar{v} - \not{v} = \gamma (\xi_m)_m \bar{v} \perp \mathcal{K}_m \end{aligned}$$

$\rightsquigarrow$

$$e_m = u_m - u_* = A^{-1} r_m = \gamma (\xi_m)_m A^{-1} \bar{v}$$

- I.e., error $e_m$ is solution of

$$A e_m = \gamma (\xi_m)_m \bar{v}$$

. . . problem of original type, with a posteriori data, and $\bar{v} = v_{m+1}$

- Extension to general rational case $\longrightarrow$

- Rational case: $\beta = \|v\|_2,\ \gamma = h_{m+1,m}$

$$u_m = V_m\,\xi_m \quad \text{with} \quad \xi_m = \beta\, R(H_m)\,\mathbf{e}_1 \in \mathbb{R}^m$$

- Let

$$e_m = u_m - u_* = V_m\,\xi_m - R(A)v$$

**Lemma.** For $m > \text{degree}(P)$

$$e_m = R_m(A)\,\bar{v} \quad \text{with} \quad R_m(z) = Q(z)^{-1} P_m(z)$$

$P_m(z)$ a posteriori defined in terms of $\xi_m$
($\text{degree}(P_m) = \text{degree}(Q) - 1$)

- For proof: Write

$$Q(z) = \sum_{j=0}^{k} q_j\, z^j$$

and evaluate residual:

$$r_m = Q(A)\,u_m - P(A)\,v = \gamma \sum_{j=0}^{k} q_j\, E_j\, \xi_m$$

with recursively defined $E_j \in \mathbb{R}^{n\times m}$, $E_1 = \bar{v}\,\mathbf{e}_m^T$

- $V := V_m,\ H := H_m$
- With $A\,V = V H + \gamma\,\bar{v}\,\mathbf{e}_m^T$, consider

$$Q(A)V = \sum_{j=0}^{k} q_j A^j V$$

- Recursion:
  - $j=0$ :

$$A^0 V = V H^0 + \gamma E_0, \quad E_0 = 0$$

  - $j=1$ :

$$A^1 V = V H^1 + \gamma E_1, \quad E_1 = \bar{v}\,\mathbf{e}_m^T$$

  - $j=2$ :

$$A^2 V = A(A^1 V) \stackrel{!}{=} A(V H^1 + \gamma E_1) = (A\,V) H^1 + \gamma A E_1$$

$$\stackrel{!}{=} (V H + \gamma E_1) H^1 + \gamma A E_1 = V H^2 + \gamma E_2, \quad E_2 = E_1 H + A E_1$$

  - $\rightsquigarrow$ For $j=0,1,2,3,\ldots$ :

$$\boxed{A^j V = V H^j + \gamma E_j, \quad E_j = E_1 H^{j-1} + A^1 E_{j-1}}$$

- Here,

$$E_3 = E_1 H^2 + A E_1 H + A^2 E_1$$

$$E_4 = E_1 H^3 + A E_1 H^2 + A^2 E_1 H + A^3 E_1, \quad \text{etc.}$$

Proof of Lemma: derive representation for residual $r_m$, (ii)

- Apply:

$$Q(A)\,u_m = Q(A)V\xi_m = \sum_{j=0}^{k} q_j\,A^j V\,\xi_m = \sum_{j=0}^{k} q_j\,(V H^j + \gamma\,E_j)\,\xi_m$$
$$= V\,Q(H)\,\xi_m + \gamma \sum_{j=0}^{k} q_j\,E_j\,\xi_m$$

- with
  - $j=0$:
  $$E_0\,\xi_m = 0$$
  - $j=1$:
  $$E_1\,\xi_m = \bar{v}\,\mathbf{e}_m^T\,\xi_m = (\mathbf{e}_m^T\,\xi_m)\,\bar{v} = (\xi_m)_m\,\bar{v}$$
  - $j=2$:
  $$E_2\,\xi_m = (E_1 H + A\,E_1)\,\xi_m = \bar{v}\,(\mathbf{e}_m^T\,H\,\xi_m)\,\bar{v} + (\xi_m)_m\,A\,\bar{v}$$
  $$= (H\,\xi_m)_m\,\bar{v} + (\xi_m)_m\,A\,\bar{v}$$
  - $j=3$:
  $$E_3\,\xi_m = \;\ldots\; = (H^2\xi_m)_m\,\bar{v} + (H\,\xi_m)_m\,A\,\bar{v} + (\xi_m)_m\,A^2\,\bar{v}$$

etc.

# Representation of approximation error

## Proof of Lemma: Representation for residual $r_m$ and error $e_m$

- $\rightsquigarrow$ residual:

$$\boxed{r_m} = V_m Q(H_m)\xi_m + \gamma \sum_{j=0}^{k} q_j E_j \xi_m - V_m P(H_m) V_m^T v$$

$$= V_m \underbrace{\left(Q(H_m)\xi_m - \beta P(H_m)\mathbf{e}_1\right)}_{=0 \text{ (FOM)}} + \gamma \sum_{j=0}^{k} q_j E_j \xi_m$$

$$= \boxed{\gamma \sum_{j=0}^{k} q_j E_j \xi_m =: P_m(A)\bar{v}}, \quad \text{degree}(P_m) = k-1$$

- $\rightsquigarrow$ error:

$$\boxed{e_m} = u_m - R(A)v = Q(A)^{-1}\left(Q(A)u_m - P(A)v\right)$$

$$= Q(A)^{-1} r_m = \boxed{R_m(A)\bar{v}, \quad R_m(z) = Q(z)^{-1} P_m(z)}$$

$\Rightarrow$ Lemma $\checkmark$

- $\rightsquigarrow$ $e_m$ = solution of problem of original type,
  with a posteriori data, and $\bar{v} = v_{m+1}$

# A posteriori error estimation

## Remarks, outlook

- Coefficients of $P_m(z)$: depend on

$$(H_m^0\,\xi_m)_m,\ (H_m^1\,\xi_m)_m,\ \ldots,\ (H_m^{k-1}\xi_m)_m, \quad k = \text{degree}(Q)$$

(optimize evaluation in course of overall process)

- For error estimation: Approximate error equation

$$e_m = R_m(A)\,\bar{v} = Q(A)^{-1}P_m(A)\,\bar{v}$$

- ... t.b.d. – consider similar techniques as for $\phi(A) = \exp(A)$
- E.g., use continued Krylov process.