The solution of large scale generalized eigenvalue problems by the rational Krylov method

Karl Meerbergen
LMS Numerical Technologies
Leuven — Belgium
km@lmsnit.be

Motivation

Aim is the solution of

$$Ax = \lambda Bx$$

Classical approach :

- Shift-invert Arnoldi
- Implicitly restarted Arnoldi

Disadvantages of Arnoldi :

- Accurate solution of linear systems is required (small backward error)
- Pole (or shift) cannot be changed unless complete restart

Rational Krylov

- Pole may change
- Backward error of solution of linear systems need not be very small
- Deflation, purging
- Implicit restart

Contents

Inexact solution of linear systems
(Joint with R. Lehoucq)

- Rational Krylov
- Inexact linear systems
- Example : Inexact Inverse Iteration
- Example : Inexact Rational Krylov
- Conclusions

Implicit application of a rational filter
(Joint with G. De Samblanx)

- Cheap inverse subspace iteration
- Example
- Conclusions

Rational Krylov : concept

Arnoldi :

Iteration k : expansion of $V_k = [v_1, ..., v_k]$ by v_{k+1}, obtained by orthogonalization of Av_k against V_k

$$\begin{aligned} Av_k &= V_{k+1} h_k \\ AV_k e_k &= V_{k+1} h_k \end{aligned}$$

For $k = 1, .., m$

$$\begin{aligned} AV_m I_m &= V_{m+1} \underline{H}_m \\ \underline{H}_m &\in \mathbb{C}^{m+1 \times m} \end{aligned}$$

Rational Krylov

Iteration k : expansion of V_k by v_{k+1}

$$(A - \mu_k B)^{-1} B V_k t_k = V_{k+1} h_k$$

μ_k : pole

t_k : continuation vector

$$AV_{k+1} h_k = BV_{k+1} \left(h_k \mu_k + \begin{bmatrix} t_k \\ 0 \end{bmatrix} \right)$$

For $k = 1, ..., m$

$$\begin{aligned} AV_{m+1} \underline{L}_m &= BV_{m+1} \underline{K}_m \\ \underline{L}_m, \underline{K}_m &\in \mathbb{C}^{m+1 \times m} \end{aligned}$$

Computation of 'Ritz' pairs in rational Krylov

Recurrence relation

$$\begin{aligned} AV_{m+1}\underline{L}_m &= BV_{m+1}\underline{K}_m \\ AVL &= BVK \end{aligned}$$

"Ritz" pairs of $B^{-1}A$ in Range(VL) :

- (θ, y) with $y = VLz$.
- Galerkin projection :

$$\begin{aligned} (VL)^*B^{-1}A(VL)z &= \theta(VL)^*(VL) \\ L^*Kz &= \theta L^*Lz \\ L^\dagger Kz &= \theta z \end{aligned}$$

- Residual is

$$\begin{aligned} B^{-1}Ay - \theta y &= V(K - \theta L)z \\ &= f \\ \|B^{-1}Ay - \theta y\| &= \|f\| = \|(K - \theta L)z\| \end{aligned}$$

Shift-invert and Cayley transform

- Shift-invert : $(A - \mu_k B)^{-1} B$
- Cayley transform : $(A - \mu_k B)^{-1}(A - \nu_k B)$

  μ_k : pole

  ν_k : zero

Use in rational Krylov

- Shift-invert : solve

$$(A - \mu_k B) w_k = B V_k t_k$$

  for w_k.

- Cayley transform : solve

$$(A - \mu_k B) w_k = (A - \nu_k B) V_k t_k$$

  for w_k.

- Special case for Cayley

  Let (θ_{k-1}, y_{k-1}) be 'Ritz' pair for iteration $k - 1$.
  Let $V_k t_k = y_{k-1}$
  Let $\nu_k = \theta_{k-1}$
  Then solve

$$(A - \mu_k B) w_k = A y_{k-1} - \theta_{k-1} B y_{k-1} =: r_{k-1}$$

Iterative solution of linear systems

Consider the linear system

$$Cx = b$$

Natural stopping criterion for linear systems solvers

$$\begin{aligned} s &= b - Cx \\ \|s\| &\leq \tau \|b\| \end{aligned}$$

where τ is the relative residual tolerance

shift-invert : $s_k = BV_k t_k - (A - \mu_k B) w_k$
$\|s_k\| \leq \tau \|B\|$

Cayley transform : $s_k = r_{k-1} - (A - \mu_k B) w_k$
$\|s_k\| \leq \tau \|r_{k-1}\|$

Inexact linear systems

Exact computation of $w_k = V_{k+1} h_k$ in

$$\begin{aligned} w_k &= (A - \mu_k B)^{-1} B V_k t_k \\ w_k &= (A - \mu_k B)^{-1} r_{k-1} \end{aligned}$$

leads to $A V_{m+1} \underline{L}_m = B V_{m+1} \underline{K}_m$

Inexact computation of w_k with residual s_k

$$\begin{aligned} s_k &= B V_k t_k - (A - \mu_k B) w_k \\ s_k &= r_{k-1} - (A - \mu_k B) w_k \end{aligned}$$

leads to

$$A V_{m+1} \underline{L}_m = B V_{m+1} \underline{K}_m - S_m$$

with $S_m = [s_1, ..., s_m]$

Alternative notation

$$\begin{aligned} AVL &= BVK - S \\ (A+E)VL &= BVK \\ E &= SL^{\dagger}V \end{aligned}$$

Ritz residual : $(\theta, y = VLz)$

$$\begin{aligned} B^{-1}Ay - \theta y &= V(K - \theta L)z - B^{-1}Sz \\ &= f - g \end{aligned}$$

Rational Krylov Algorithm

1. Given v_1, $\|v_1\| = 1$.
   Let $y_0 = v_1$ and $\theta_0 = 0$
2. For $k = 1, 2, ..., m$
   - 2.1. Compute $r_{k-1} = Ay_{k-1} - \theta_{k-1} By_{k-1}$
   - 2.2. If $\|r_{k-1}\| <$ TOL then exit
   - 2.3. Select $\mu_k \neq \theta_{k-1}$.
   - 2.4. Solve $(A - \mu_k B)w = r_{k-1}$.
   - 2.5. Orthonormalize w against V_k into v_{k+1}.
   - 2.6. Update $\underline{L}_k$ and $\underline{K}_k$.
   - 2.7. Solve $\underline{L}_k^{\dagger} \underline{K}_k z = \theta z$
   - 2.8. Select 'Ritz' value θ_k and
   - 2.9. 'Ritz' vector $y_k = V_{k+1} \underline{L}_k z_k / \|\underline{L}_k z_k\|$.

Exact inverse iteration

Special case of Rational Krylov :

Let $y_k = V_{k+1}\underline{L}_k z_k$ with $z_k = e_k/\|\underline{L}_k e_k\|$

Then $\{y_k\}_{k=1,\ldots,m}$ is the inverse iteration sequence

$$\begin{aligned} y_0 &= v_1 \\ y_1 &= V_2 z_2 = \zeta_1 (A-\mu_1 B)^{-1} B y_0 \\ y_2 &= V_3 z_3 = \zeta_2 (A-\mu_2 B)^{-1} B y_1 \\ &\vdots \end{aligned}$$

Rayleigh quotient :

$$\theta_k = y_k^* B^{-1} A y_k = \frac{e_k^* \underline{L}_k^* \underline{K} e_k}{\|\underline{L}_k e_k\|^2}$$

Residual for exact inverse iteration :

$r_k = B^{-1} y_k - \theta_k y_k = f_k$

Assume that $\|f_k\| \leq \rho^k \|f_0\|$

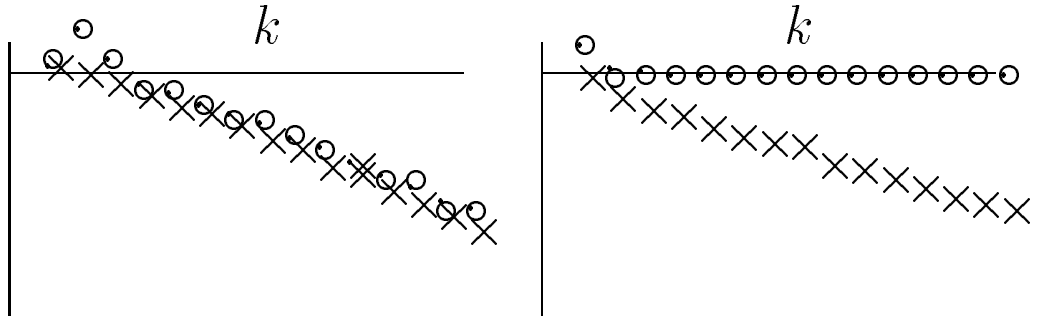
Inexact inverse iteration

Analysis for $B = I$:

With Cayley transform	With shift-invert
$\|f_k\| \leq \rho^k \|f_0\|$	$\|f_k\| \leq \rho^k \|f_0\|$
$\|s_k\| \leq \tau \|r_{j-1}\|$	$\|s_k\| \leq \tau$
$\|r_k\| \leq (\rho + \tau)^k \|f_0\| + \rho^k \|r_0\|$	$\|r_k\| \leq \rho^k \|f_0\| + \tau$

Example : (Olmstead model — $n = 100$)
$\mu_k = 5$ — Gauss-Seidel relaxation

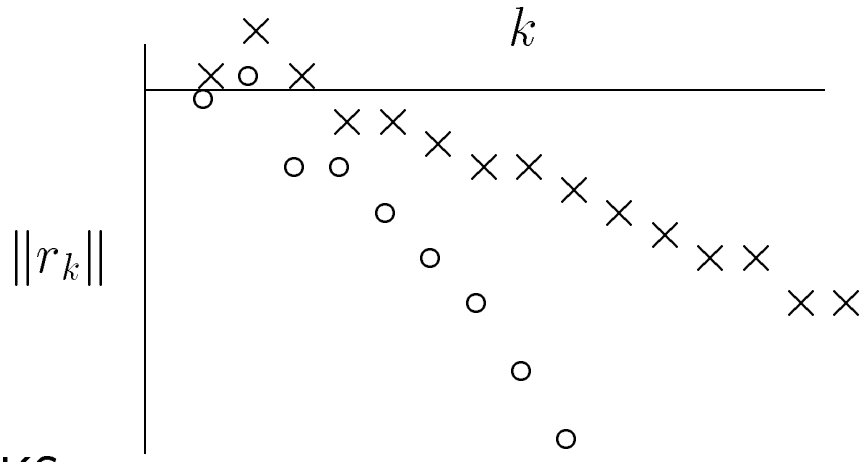
Cayley : Shift-invert :



$\|f_k\|$: $\times$
$\|s_k\|$: $\cdot$
$\|r_k\|$: $\circ$

Inexact Rational Krylov

Example :



$\circ$: RKS

$\times$: inverse iteration

Tilted plane problem

Simulation of viscous free surface flow on a tilted plane

$$Ax = \lambda Bx \qquad B \text{ singular} \qquad n = 536$$

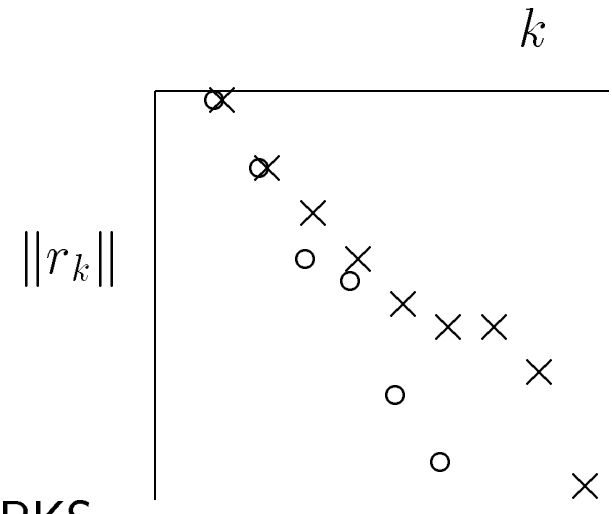
We solve instead

$$Bx = \gamma Ax \qquad \gamma = \lambda^{-1}$$

RKS with fixed pole $\mu_k = -0.1$

Linear systems are solved with Sparskit routine PGMRES using ILUT(lfil=40,tol=1.e-3) and relative residual tolerance $\tau = 10^{-4}$

Comparison of RKS and Inverse Iteration



$\circ$: RKS

$\times$: inverse iteration

Jacobi-Davidson and Rational Krylov

Let y_{k-1} be $k-1$st 'Ritz' vector and

$$\theta_{k-1} = \frac{y_{k-1}^* A y_{k-1}}{y_{k-1}^* B y_{k-1}}$$

Solve the linear system

$$(I - \frac{Byy^*}{y^*By})(A - \mu_k B)(I - \frac{yy^*}{y^*y})w = r_{k-1}$$

A solution of this system is given by the solution of the system

$$\begin{aligned}(A - \mu_k B)w &= \epsilon_k B y_{k-1} + r_{k-1} \\ &= (A - (\theta_{k-1} + \epsilon_k)B)y_{k-1}\end{aligned}$$

which represents a specific Cayley transformation

Moreover,

$$\epsilon_k = \frac{y_{k-1}^*(A - \mu_k B)w}{y_{k-1}^* B y_{k-1}}$$

Conclusions

Advantages

- Recurrence matrices $\underline{L}_m$ and $\underline{K}_m$ are cheaply computed.
- No explicit formation of V^*AV and V^*BV required
- RKS tools are (probably) applicable
  - Use of real arithmetic when poles are complex
  - Harmonic Ritz values
  - deflation, purging and implicit filtering.

Disadvantages ??

- No Galerkin projection method
- It is tricky to include Newton scheme (JD)

Implicit rational filter

Original RKS relation

$$AV_{k+1}\underline{L}_k = BV_{k+1}\underline{K}_k$$

Transformation to a new basis (with Q_{k+1} square and unitary)

$$V_{k+1}^{+} = V_{k+1}Q_{k+1}$$

New RKS

$$AV_{k+1}^{+}\underline{L}_k^{+} = BV_{k+1}^{+}\underline{K}_k^{+}$$

with

$$\underline{L}_k^{+} = Q_{k+1}^{*}\underline{L}_k Z_k$$
$$\underline{K}_k^{+} = Q_{k+1}^{*}\underline{K}_k Z_k$$

where Z_k is square and unitary

Implicit filter

Choose α and β such that $\alpha\mu_j \neq \beta$ for $j = 1, ..., k$.

Let Q_{k+1} be chosen such that

$$\alpha \underline{K}_k - \beta \underline{L}_k = Q_{k+1} \underline{R}_k$$

($\underline{R}$ is upper triangular and Q is unitary upper Hessenberg)

Find an upper Hessenberg Z_k such that

$$Q_{k+1}^* (\bar{\beta} \underline{K}_k + \bar{\alpha} \underline{L}_k) Z_k$$

is upper Hessenberg, then

$$\underline{K}_k^+ \quad \text{and} \quad \underline{L}_k^+$$

are upper Hessenberg and $\underline{L}_k^+$ is unreduced.
In practice it is sufficient that Z_k maps the last row of

$$Q_{k+1}^* (\bar{\beta} \underline{K}_k + \bar{\alpha} \underline{L}_k)$$

into $\begin{bmatrix} 0 & \cdots & 0 & * \end{bmatrix}$.

Implicit filter (cont...)

After the transformation

$$AV_{k+1}^{+}\underline{L}_{k}^{+} = BV_{k+1}^{+}\underline{K}_{k}^{+}$$

reduce the order by cutting the last column :

$$AV_{k}^{+}\underline{L}_{k-1}^{+} = BV_{k}^{+}\underline{K}_{k-1}^{+}$$

Properties

- Range$(V_k^+) = (A - \mu_k B)^{-1}(\alpha A - \beta B)$Range$(V_k)$
- When β/α is a Ritz value (exact shift), then this corresponds to a purging of the corresponding Schur form.

Example

Simulation of viscous free surface flow on a tilted plane

$$Ax = \lambda Bx \qquad B \text{ singular}$$

Infinite eigenvalue can disturb the computation.
The eigenspace is the null-space of $(A - \mu B)^{-1}B$ for any μ.

Corruption by the infinite eigenvalue can be avoided by removing the eigenspace from the Krylov subspace by

$$V^+ = (A - \mu B)^{-1}BV$$

This is achieved in a cheap way by a filter with $\alpha = 0$ and $\beta = -1$.

We performed RKS with Gaussian elimination
The pole was set to the rightmost harmonic Ritz value after the 3rd, 6th and 9th iterations.

The poles on the 1st, 4th, 7th and 10th iterations

iteration	with filter	without filter
1	0.0	0.0
4	−8.4677	8.43
7	−9.4883	19.75
10	−9.4883	74.83

Conclusions

- Implicit application of subspace iteration (cheap)
- Properties similar to implicitly restarted Arnoldi
- Open questions
  - How to purge/deflate ?
  - Numerical stability of implicit filter
  - How to combine with Inexact Rational Krylov

References

- R. Lehoucq and K. Meerbergen. The inexact Rational Krylov Sequence Method. Preprint 978. Department of Mathematics. Utrecht University. 1996. Submitted for publication.
- K. Meerbergen and D. Roose. The restarted Arnoldi method applied to iterative linear system solvers for the computation of rightmost eigenvalues. SIAM J. Matrix Anal. Appl. 18, pp.1–20, 1997.
- G. De Samblanx, K. Meerbergen and A. Bultheel. The implicit application of a rational filter in the RKS method. Technical Report TW239, Department of Computer Science, K.U.Leuven. 1996. Accepted for publication in BIT.