On the implicit restart of the rational Krylov method

Chasing algorithms for polynomial, extended and rational Krylov

Daan Camps, Karl Meerbergen, and Raf Vandebril

July 27, 2017

KU Leuven - University of Leuven - Department of Computer Science - NUMA

Problem statement

Problem setting

Single sentence problem statement

We want to compute a small number of eigenpairs of $A \in \mathbb{C}^{N \times N}$ that satisfy some property $\mathfrak{P}$.

Examples:

- All eigenvalues in a domain $\Omega \subset \mathbb{C}$
- The eigenvalue(s) with largest real part
- The eigenvalue(s) closest to the imaginary axis

Polynomial Krylov

Polynomial Krylov

Definition (Krylov subspace)

Given a matrix $A \in \mathbb{C}^{N \times N}$ and a vector $\mathbf{0} \neq \boldsymbol{v} \in \mathbb{C}^{N}$:

$$\mathcal{K}_{m+1} = \mathcal{K}_{m+1}(A, \boldsymbol{v}) := \operatorname{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^m \boldsymbol{v}\}.$$

Polynomial Krylov

Definition (Krylov subspace)

Given a matrix $A \in \mathbb{C}^{N\times N}$ and a vector $\mathbf{0} \neq \boldsymbol{v} \in \mathbb{C}^N$:

$$\mathcal{K}_{m+1} = \mathcal{K}_{m+1}(A, \boldsymbol{v}) := \text{span}\{\boldsymbol{v}, A\boldsymbol{v}, \dots, A^m\boldsymbol{v}\}.$$

- Assumption: subspace is A-variant: $A\,\mathcal{K}_{m+1} \not\subseteq \mathcal{K}_{m+1}$.
- Isomorphism: $\mathcal{K}_{m+1} \cong \mathcal{P}_m$, i.e. $\forall \boldsymbol{w} \in \mathcal{K}_{m+1}, \exists p \in \mathcal{P}_m : \boldsymbol{w} = p(A)\boldsymbol{v}$ and vice versa.
- Orthogonal basis $V \in \mathbb{C}^{N\times(m+1)}$ such that

$$\text{span}\{V_{(:,1:k+1)}\} = \text{span}\{\boldsymbol{v}, A\boldsymbol{v}, \dots, A^k\boldsymbol{v}\} \qquad \forall k \leq m.$$

Polynomial Krylov

Arnoldi's method: recurrence relation

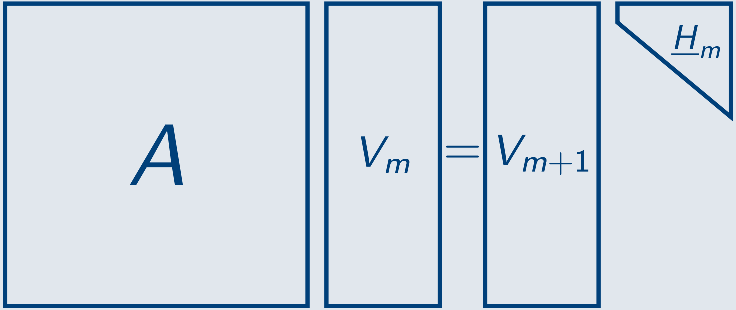
$$A\,V_m = V_m\,H_m + \,h_{m+1,m}\,v_{m+1}e_m^T$$

Polynomial Krylov

Arnoldi's method: recurrence relation

$$A\,V_m = V_{m+1}\,\underline{H}_m$$

Arnoldi's method: recurrence relation



Polynomial Krylov

How to extract eigenpairs from $\mathcal{K}_{m+1}$?

Polynomial Krylov

How to extract eigenpairs from $\mathcal{K}_{m+1}$?

$\Rightarrow$ Compute the **Ritz pairs**

Small scale eigenvalue problem:

$$H_m \boldsymbol{z} = \vartheta \boldsymbol{z}$$

Ritz pairs $(\vartheta, \boldsymbol{x}) := (\vartheta, V_m \boldsymbol{z})$ satisfy Galerkin condition $A\boldsymbol{x} - \vartheta \boldsymbol{x} \perp \mathcal{K}_m$

Polynomial Krylov

What if the eigenvalues of A that satisfy $\mathfrak{P}$ only converge for $m \to N$?

Polynomial Krylov

What if the eigenvalues of A that satisfy $\mathfrak{P}$ only converge for $m \to N$?

- **Explicit restart:** for certain maximal m, select $\boldsymbol{w} \in \mathcal{K}_{m+1}$ and continue constructing new $\mathcal{K}_{m+1}(A, \boldsymbol{w})$ from scratch (Saad, 1980).
- **Implicit restart:** for certain maximal m, apply l-th order *polynomial filter* and continue from k-th order $\mathcal{K}_{k+1}(A, \hat{\boldsymbol{v}})$ $(l + k = m)$ (Sorensen, 1992).
- **Krylov-Schur:** compute and reorder the Schur decomposition of H_m (Stewart, 2001).

Polynomial Krylov

Implicitly restarted Arnoldi (IRA)

Input: $(V_{m+1}, \underline{H}_m), \{\varrho_i\}_{i=1}^{l}$

1: **for** $i = 1 \dots l$ **do**

2: $\quad \underline{H}_{m-i+1} - \varrho_i \underline{I}_{m-i+1} = \begin{bmatrix} Q & q \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix}$

3: $\quad \underline{H}_{m-i} := Q^* \underline{H}_{m-i+1} Q_{(1:m-i+1,1:m-i)}$

4: $\quad V_{m-i+1} := V_{m-i+2} Q$

5: **end for**

6: **return** $\hat{V}_{k+1} := V_{k+1}, \hat{\underline{H}}_k := \underline{H}_k$

Polynomial Krylov

Implicitly restarted Arnoldi (IRA)

Input: $(V_{m+1}, \underline{H}_m), \{\varrho_i\}_{i=1}^{l}$

1: **for** $i = 1 \ldots l$ **do**

2: $\underline{H}_{m-i+1} - \varrho_i \underline{I}_{m-i+1} = \begin{bmatrix} Q & q \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix}$

3: $\underline{H}_{m-i} := Q^* \underline{H}_{m-i+1} Q_{(1:m-i+1,1:m-i)}$

4: $V_{m-i+1} := V_{m-i+2} Q$

5: **end for**

6: **return** $\hat{V}_{k+1} := V_{k+1}, \underline{\hat{H}}_k := \underline{H}_k$

Result: $\hat{V}_{k+1} \boldsymbol{e}_1 = \prod_{i=1}^{l} (A - \varrho_i I) \boldsymbol{v}$

Polynomial Krylov

Practical implementation: Implicit-Q theorem

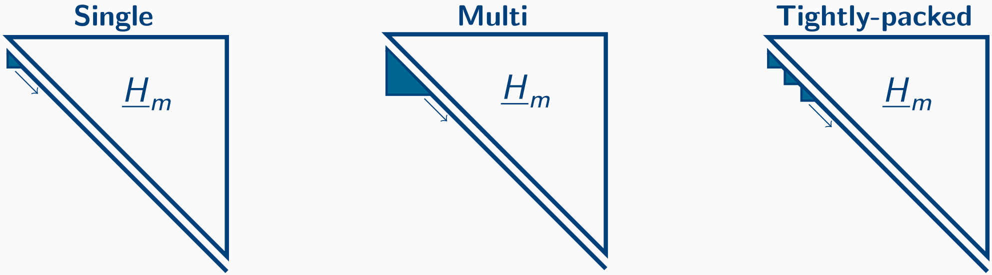
The Arnoldi decomposition $(V_{m+1}, \underline{H}_m)$ of order m is uniquely[1] defined by the matrix A and first column $V\boldsymbol{e}_1$.

[1] *essential uniqueness*

Polynomial Krylov

Practical implementation: Implicit-Q theorem

The Arnoldi decomposition $(V_{m+1}, \underline{H}_m)$ of order m is uniquely[1] defined by the matrix A and first column $V\boldsymbol{e}_1$.



[1] *essential uniqueness*

Polynomial Krylov

Illustrative example

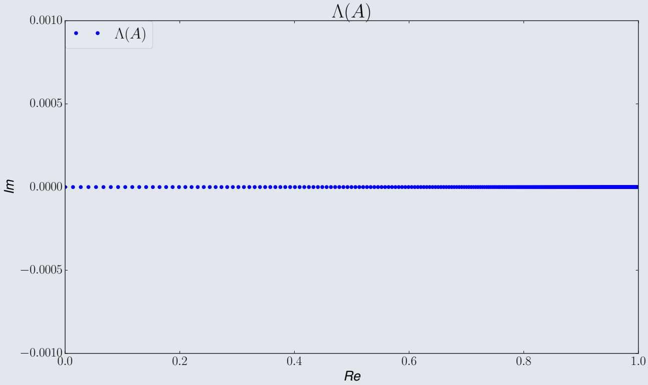
$A \in \mathbb{R}^{500\times 500}$, $\Lambda(A) \subset [0,1]$, $\mathfrak{P}$: rightmost eigenvalue

Diagonal matrix, logarithmically spaced eigenvalues

Illustrative example

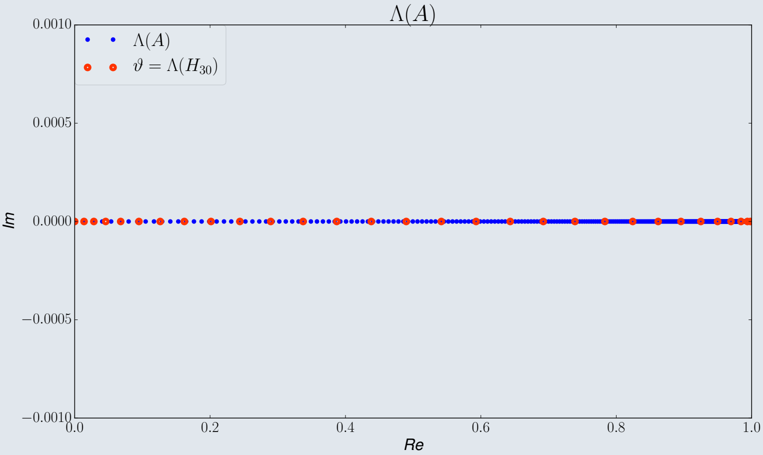
$A \in \mathbb{R}^{500\times 500}$, $\Lambda(A) \subset [0,1]$, $\mathfrak{P}$: rightmost eigenvalue

Diagonal matrix, logarithmically spaced eigenvalues



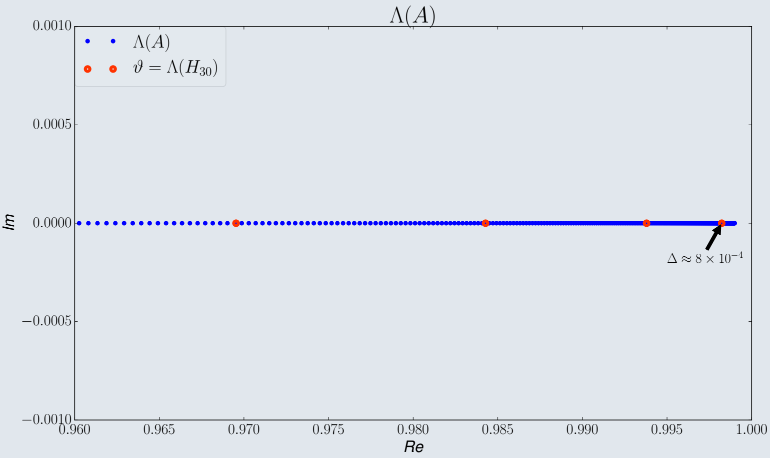
Polynomial Krylov

Illustrative example



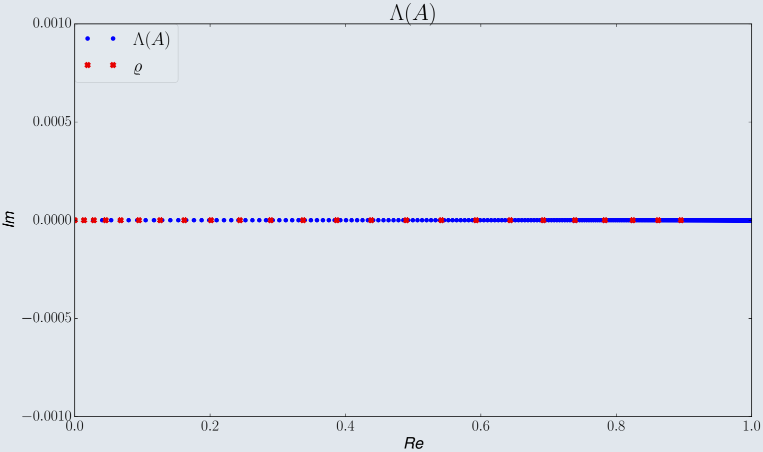
Polynomial Krylov

Illustrative example



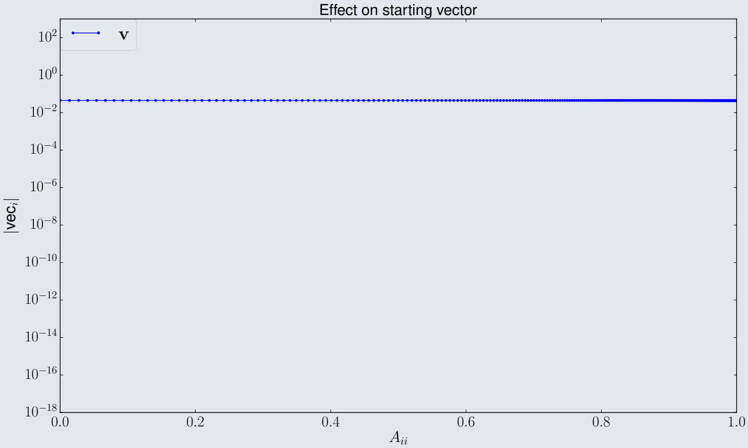
Polynomial Krylov

Illustrative example



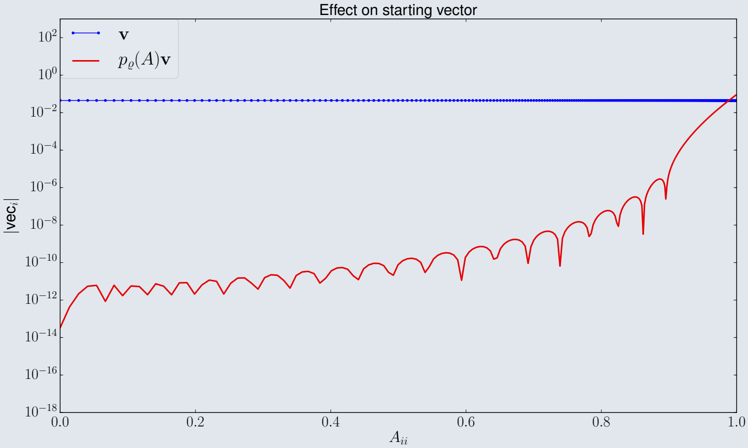
Polynomial Krylov

Illustrative example



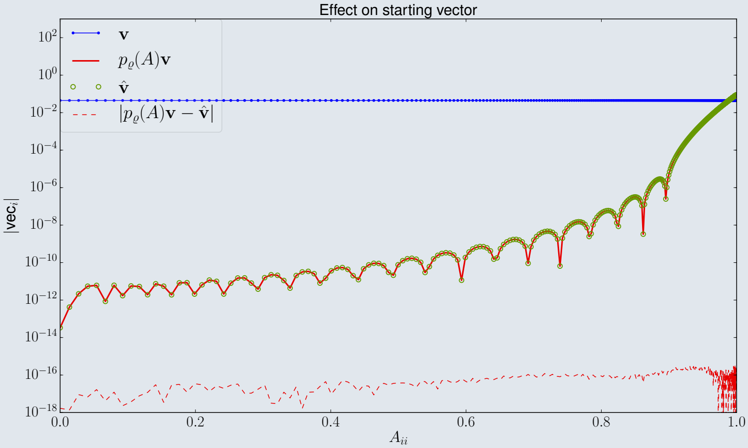
Polynomial Krylov

Illustrative example



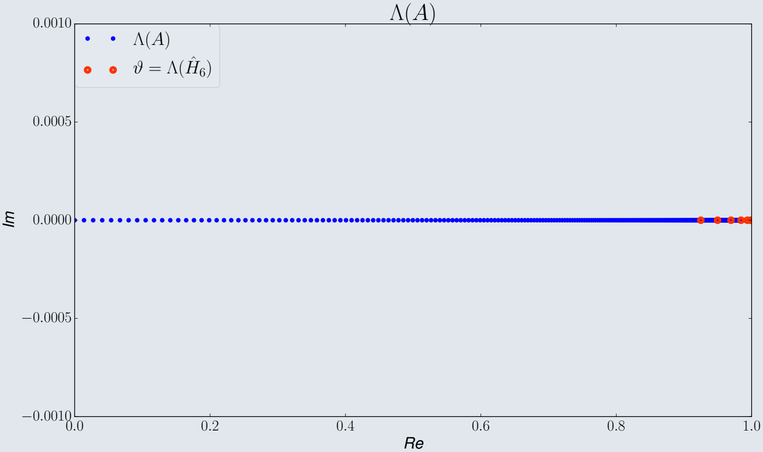
Polynomial Krylov

Illustrative example



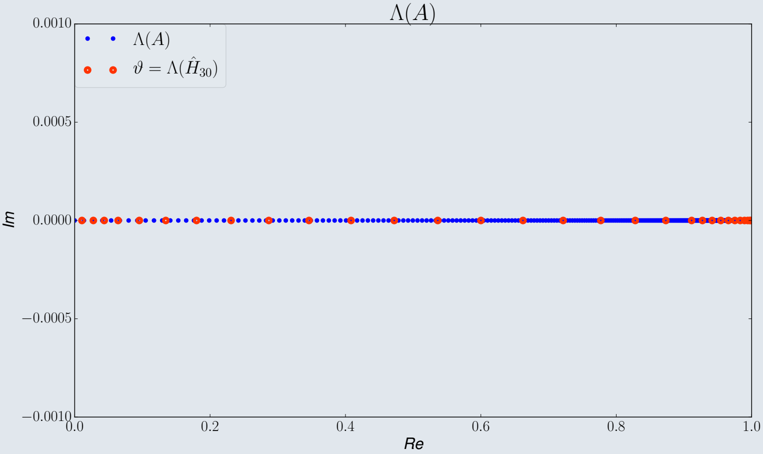
Polynomial Krylov

Illustrative example



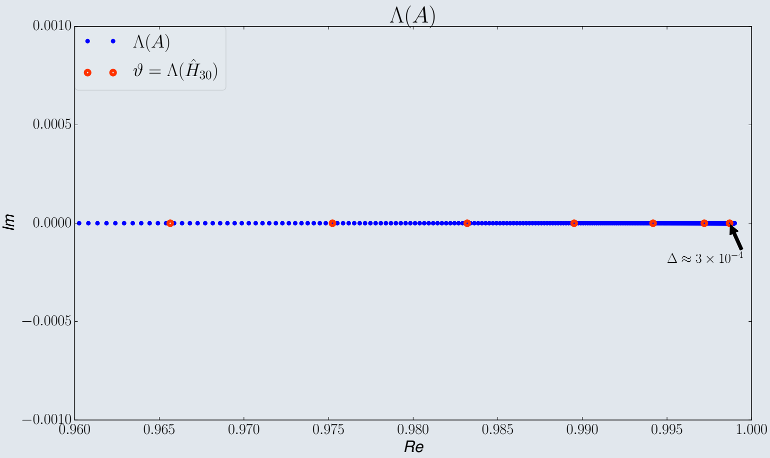
Polynomial Krylov

Illustrative example



Polynomial Krylov

Illustrative example



Rational Krylov

Rational Krylov

Definition (Rational Krylov sequence (Berljafa and Güttel, 2015))

Given a matrix $A \in \mathbb{C}^{N\times N}$, a vector $\mathbf{0} \neq \boldsymbol{v} \in \mathbb{C}^N$ and $q_m \in \mathcal{P}_m$ with roots $\Xi = \{\xi_1, \ldots, \xi_m\} \in \overline{\mathbb{C}} \setminus \Lambda(A)$:

$$\mathcal{Q}_{m+1}(A, \boldsymbol{v}, \Xi) = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) := q_m(A)^{-1}\,\mathcal{K}_{m+1}(A, \boldsymbol{v}).$$

Rational Krylov

Definition (Rational Krylov sequence (Berljafa and Güttel, 2015))

Given a matrix $A \in \mathbb{C}^{N\times N}$, a vector $\mathbf{0} \neq \boldsymbol{v} \in \mathbb{C}^N$ and $q_m \in \mathcal{P}_m$ with roots $\Xi = \{\xi_1, \ldots, \xi_m\} \in \overline{\mathbb{C}} \setminus \Lambda(A)$:

$$\mathcal{Q}_{m+1}(A, \boldsymbol{v}, \Xi) = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) := q_m(A)^{-1}\,\mathcal{K}_{m+1}(A, \boldsymbol{v}).$$

Isomorphism: $\mathcal{Q}_{m+1} \cong q_m(z)^{-1}\mathcal{P}_m$, i.e. $\forall \boldsymbol{w} \in \mathcal{Q}_{m+1}, \exists \boldsymbol{p} \in \mathcal{P}_m : \boldsymbol{w} = q_m(A)^{-1}p(A)\boldsymbol{v}$

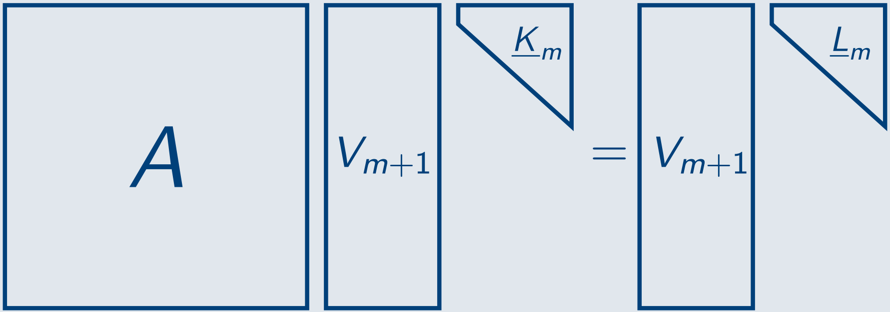
Rational Krylov

Rational Arnoldi's method (Ruhe, 1998): recurrence relation

$$A\,V_{m+1}\,\underline{K}_m = V_{m+1}\,\underline{L}_m$$

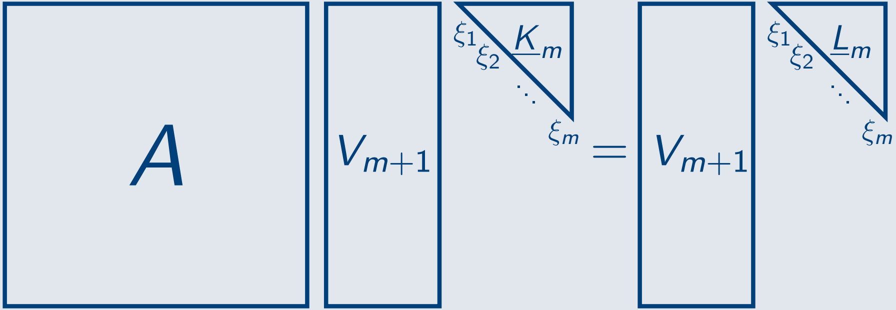
Rational Krylov

Rational Arnoldi's method (Ruhe, 1998): recurrence relation



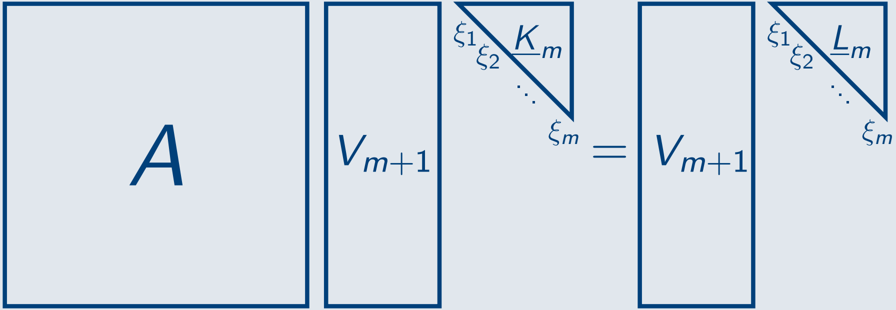
Rational Krylov

Rational Arnoldi's method (Ruhe, 1998): recurrence relation



Rational Krylov

Rational Arnoldi's method (Ruhe, 1998): recurrence relation



*(Abuse of) notation: $l_{i+1,i}/k_{i+1,i} = \xi_i$

Rational Krylov

How to extract eigenpairs from $\mathcal{Q}_{m+1}$?

Rational Krylov

How to extract eigenpairs from $\mathcal{Q}_{m+1}$?

$\Rightarrow$ Compute the **Ritz pairs**

Small scale eigenvalue problem:

$$\underline{K}_m^{\dagger} \underline{L}_m \, \boldsymbol{z} = \vartheta \, \boldsymbol{z}$$

Ritz pairs $(\vartheta, \boldsymbol{x}) := (\vartheta, V_{m+1} \underline{K}_m \, \boldsymbol{z})$ satisfy Galerkin condition
$A \, \boldsymbol{x} - \vartheta \, \boldsymbol{x} \perp \mathcal{K}_m(A, q_m(A)^{-1} \boldsymbol{v})$

Rational Krylov

What if the eigenvalues of A that satisfy $\mathfrak{P}$ only converge for $m \to N$?

Rational Krylov

What if the eigenvalues of A that satisfy $\mathfrak{P}$ only converge for $m \to N$?

Same solutions as before!

- **Explicit restart:** for certain maximal m, select $\boldsymbol{w} \in \mathcal{Q}_{m+1}$ and continue constructing new $\mathcal{Q}_{m+1}(A, \boldsymbol{w}, \hat{q}_m)$ from scratch.
- **Implicit restart:** for certain maximal m, apply l-th order *polynomial filter* and continue from k-th order $\mathcal{Q}_{k+1}(A, \hat{\boldsymbol{v}}, q_k)$ $(l + k = m)$ (De Samblanx et al., 1997).
- **Krylov-Schur:** compute and reorder the generalized Schur decomposition of (L_m, K_m) (De Samblanx et al., 1997).

Rational Krylov

How to apply the polynomial filter implicitly?

Rational Krylov

How to apply the polynomial filter implicitly?

Special case: extended Krylov introduced by Druskin and Knizhnerman (1998)

$$\Xi_{\text{ext}} = \{\xi_i\}_{i=1}^{m}, \forall i : \xi_i \in \{0, \infty\}$$

Rational Krylov

How to apply the polynomial filter implicitly?

Special case: extended Krylov introduced by Druskin and Knizhnerman (1998)

$$\Xi_{\text{ext}} = \{\xi_i\}_{i=1}^m, \; \forall i : \xi_i \in \{0, \infty\}$$

Example: $\Xi_{\text{ext}} = \{0, \infty, \ldots, 0\}$

The matrix pencil is in *condensed* format, chasing by elementary unitary operations (Camps et al., 2016)

Rational Krylov

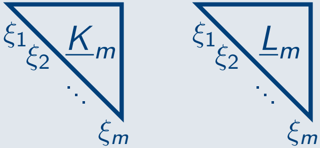
How to apply the polynomial filter implicitly?

General case:

No longer in condensed format!

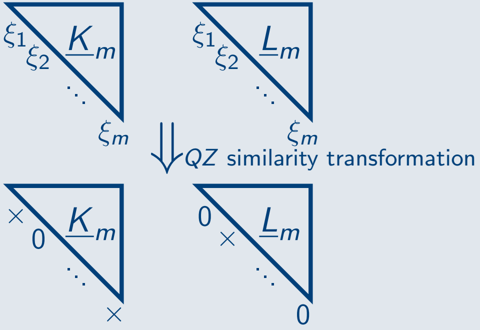
Rational Krylov

Solution



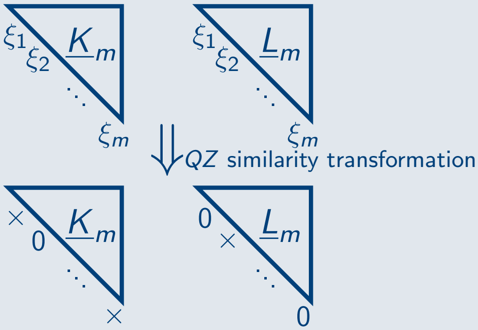
Rational Krylov

Solution



Rational Krylov

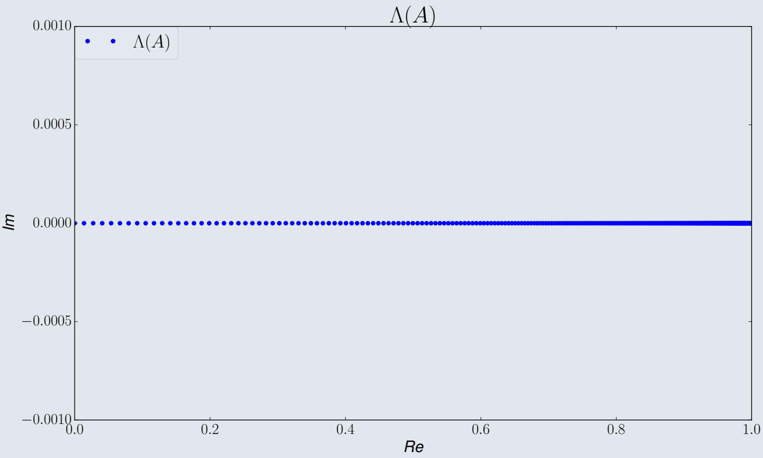
Solution



The QZ transformation is formulated in terms of elementary transformations on core transformations

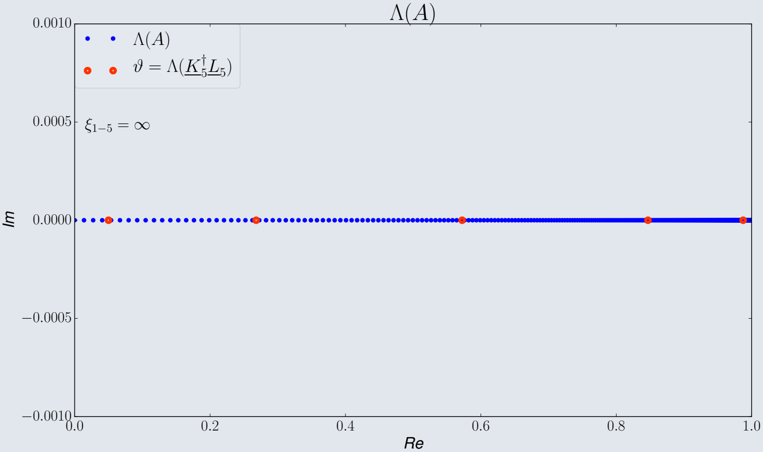
Rational Krylov

Illustrative example



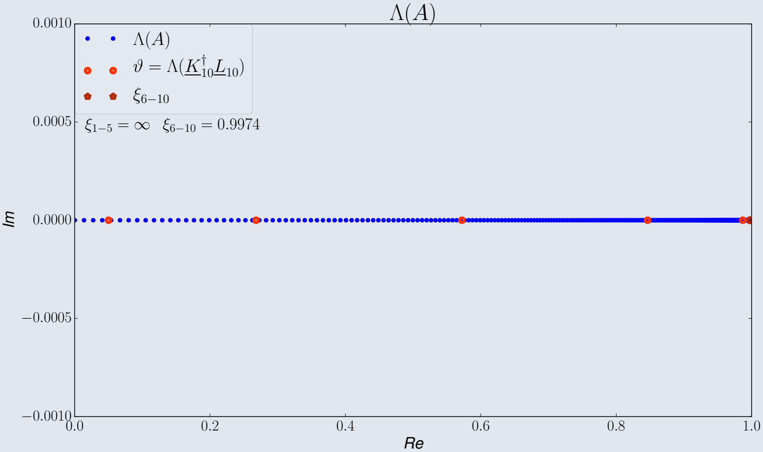
Rational Krylov

Illustrative example



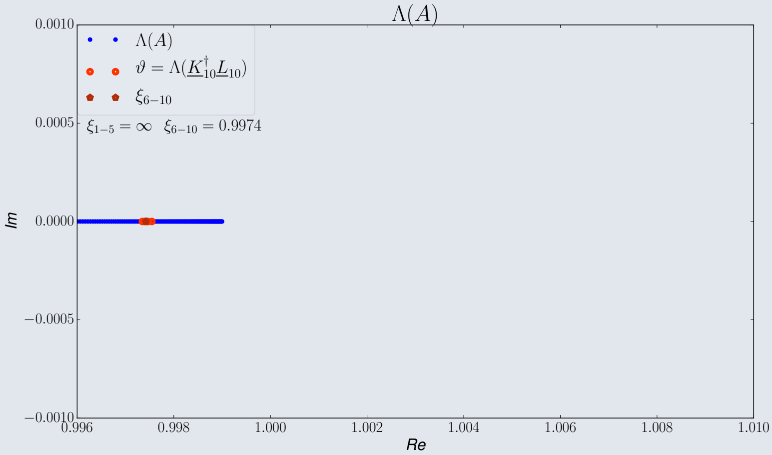
Rational Krylov

Illustrative example

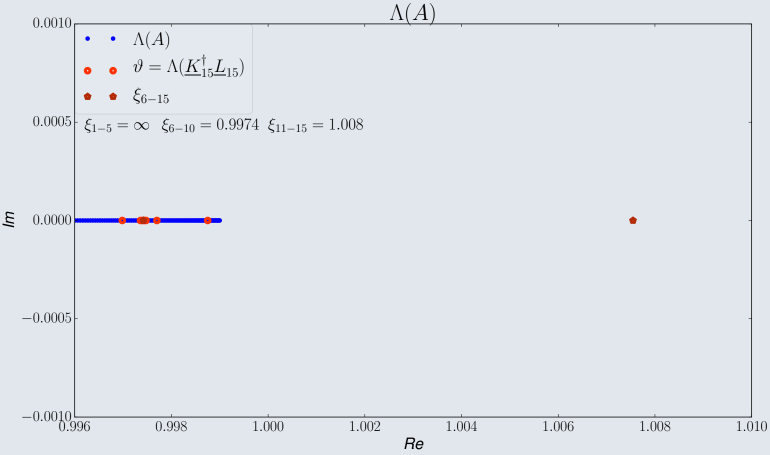


Rational Krylov

Illustrative example

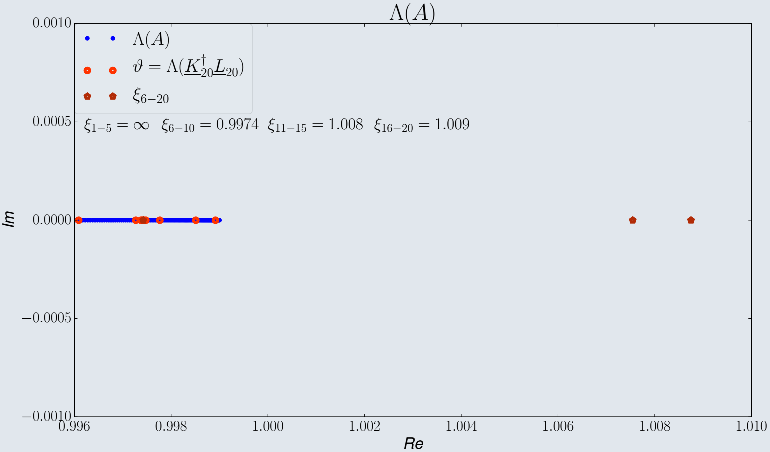


Illustrative example



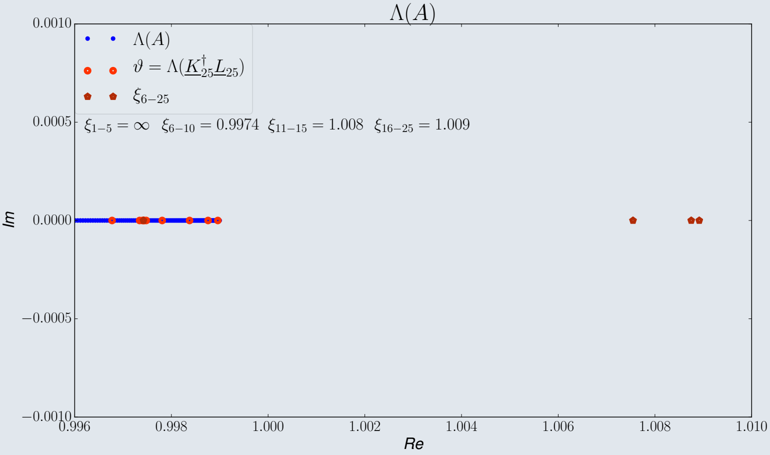
Rational Krylov

Illustrative example



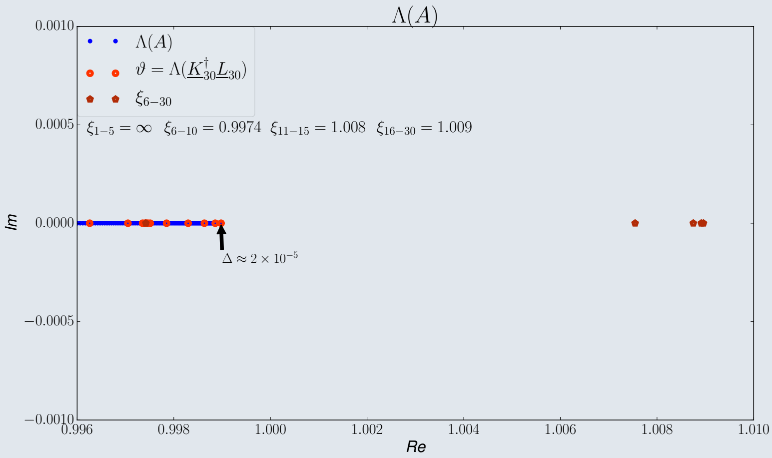
Rational Krylov

Illustrative example



Rational Krylov

Illustrative example



Rational Krylov

Illustrative example

Transformation: $(V_{m+1}, \underline{K}_m, \underline{L}_m) \rightarrow (\tilde{V}_{m+1}, \underline{\tilde{K}}_m, \underline{\tilde{L}}_m)$ with $\Xi_{\text{ext}} = \{\infty, 0, \infty, 0, \ldots, 0\}$

Rational Krylov

Illustrative example

Transformation: $(V_{m+1}, \underline{K}_m, \underline{L}_m) \to (\tilde{V}_{m+1}, \underline{\tilde{K}}_m, \underline{\tilde{L}}_m)$ with $\Xi_{\text{ext}} = \{\infty, 0, \infty, 0, \dots, 0\}$

$$\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) = \tilde{\mathcal{Q}}_{m+1}(A, \boldsymbol{w}, \Xi_{\text{ext}}) \quad \Rightarrow \quad \boldsymbol{w} = \alpha\, q_m(A)^{-1} A^{m/2}\, \boldsymbol{v}$$

Rational Krylov

Illustrative example

Transformation: $(V_{m+1}, \underline{K}_m, \underline{L}_m) \to (\tilde{V}_{m+1}, \underline{\tilde{K}}_m, \underline{\tilde{L}}_m)$ with $\Xi_{\text{ext}} = \{\infty, 0, \infty, 0, \ldots, 0\}$

$$\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) = \tilde{\mathcal{Q}}_{m+1}(A, \boldsymbol{w}, \Xi_{\text{ext}}) \quad \Rightarrow \quad \boldsymbol{w} = \alpha\, q_m(A)^{-1} A^{m/2} \boldsymbol{v}$$

Ritz values are **not** affected: $\underline{\tilde{K}}_m = Q\, \underline{K}_m\, Z, \quad \underline{\tilde{L}}_m = Q\, \underline{L}_m\, Z$

Rational Krylov

Illustrative example

Transformation: $(V_{m+1}, \underline{K}_m, \underline{L}_m) \to (\tilde{V}_{m+1}, \underline{\tilde{K}}_m, \underline{\tilde{L}}_m)$ with $\Xi_{\text{ext}} = \{\infty, 0, \infty, 0, \ldots, 0\}$

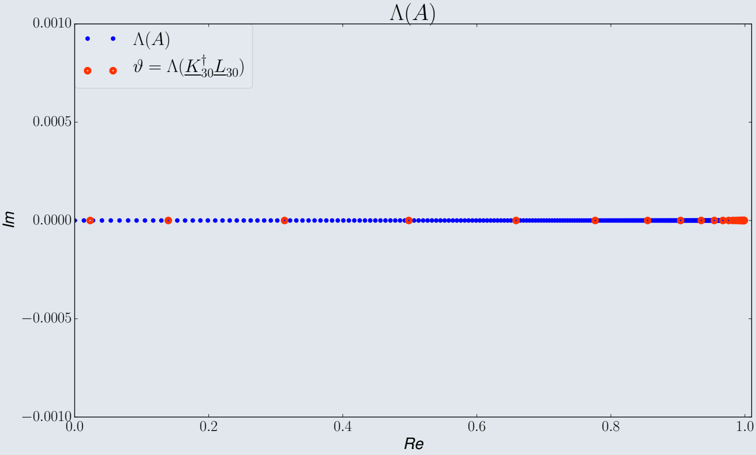
$$\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) = \tilde{\mathcal{Q}}_{m+1}(A, \boldsymbol{w}, \Xi_{\text{ext}}) \quad \Rightarrow \quad \boldsymbol{w} = \alpha\, q_m(A)^{-1} A^{m/2}\, \boldsymbol{v}$$

Ritz values are **not** affected: $\underline{\tilde{K}}_m = Q\, \underline{K}_m\, Z, \quad \underline{\tilde{L}}_m = Q\, \underline{L}_m\, Z$

$$\underline{\tilde{K}}_m^{\dagger} \underline{\tilde{L}}_m = Z^* \underline{K}_m^{\dagger} Q^* Q \underline{L}_m Z$$

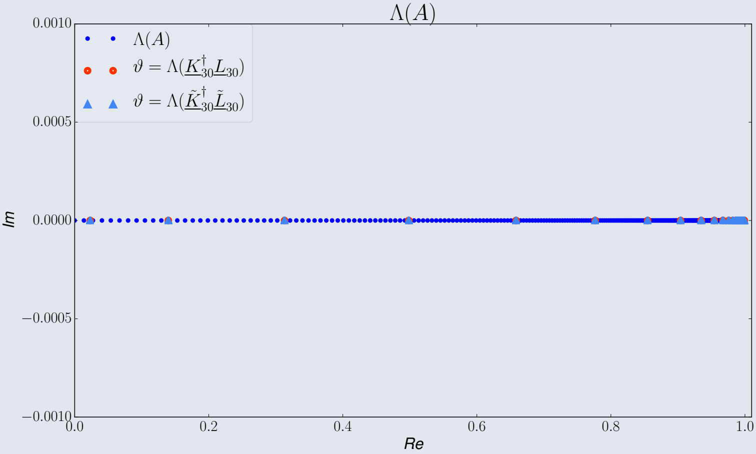
Rational Krylov

Illustrative example



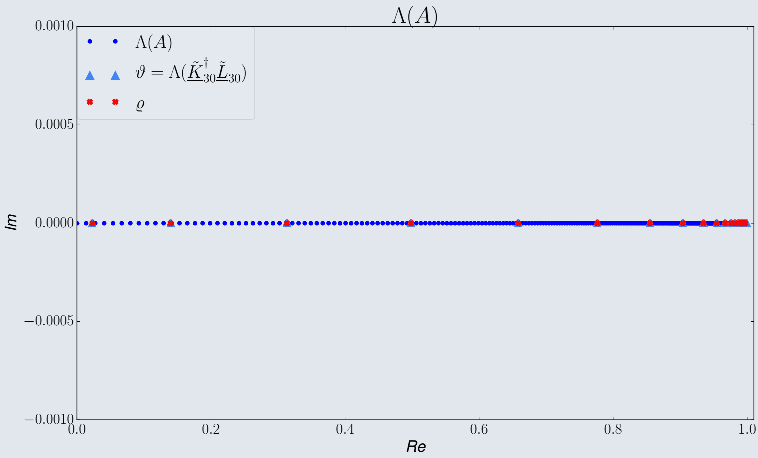
Rational Krylov

Illustrative example

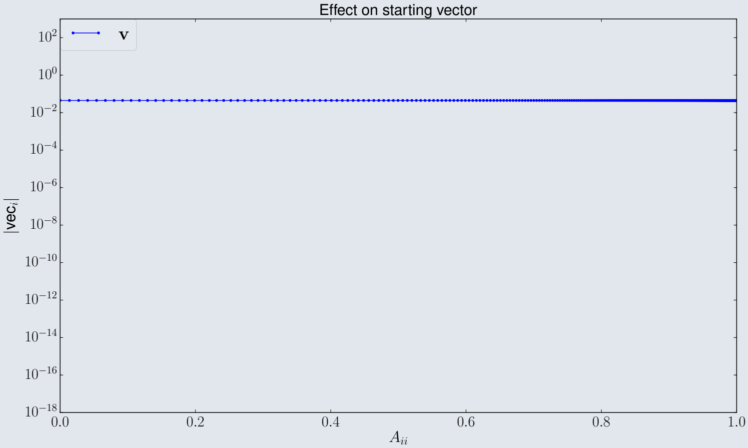


Rational Krylov

Illustrative example

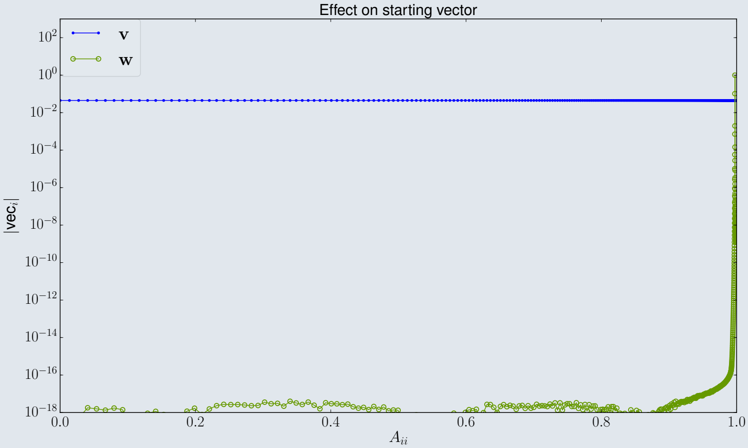


Illustrative example



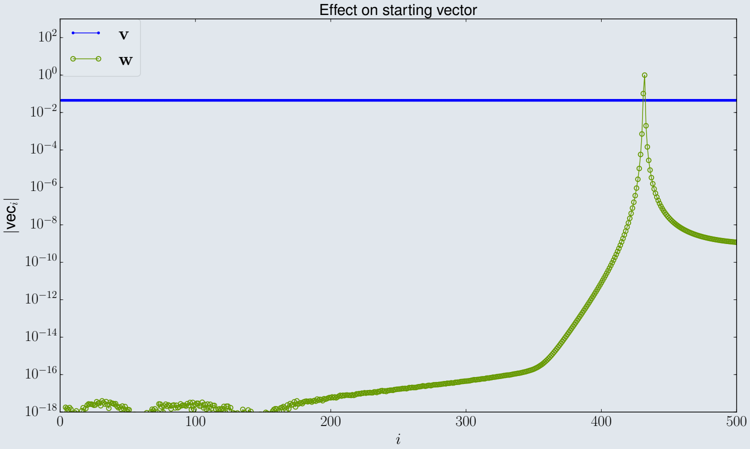
Rational Krylov

Illustrative example



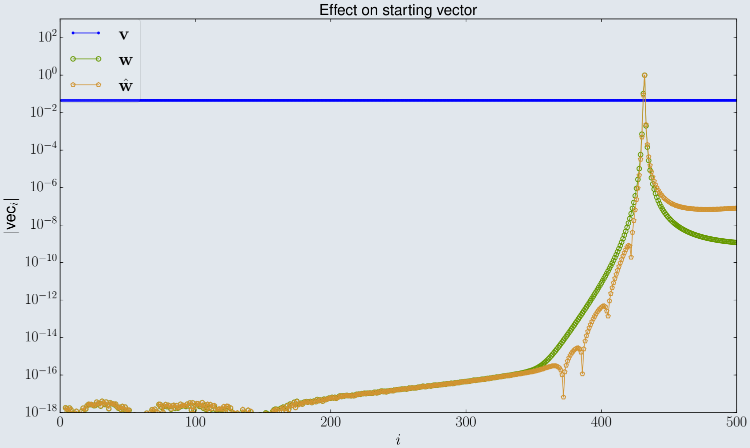
Rational Krylov

Illustrative example



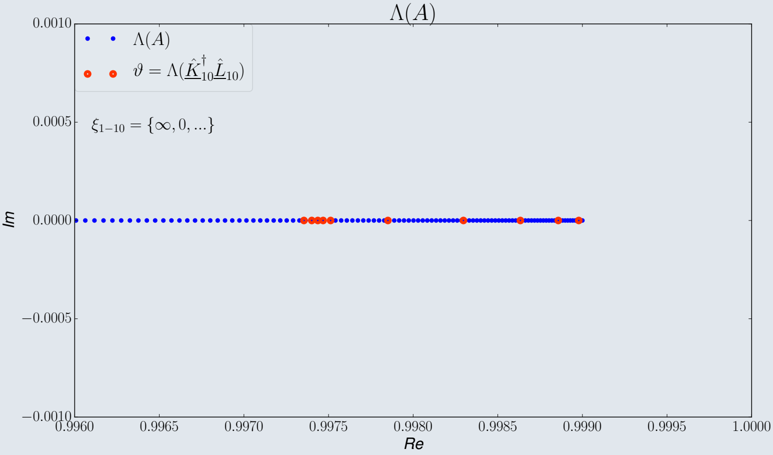
Rational Krylov

Illustrative example



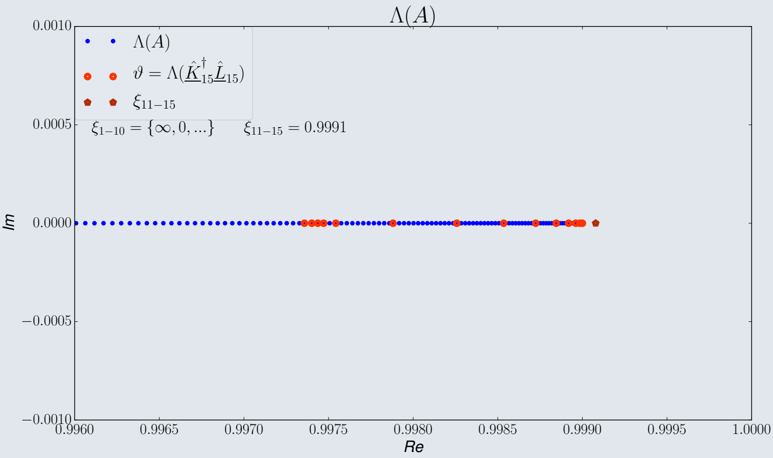
Rational Krylov

Illustrative example

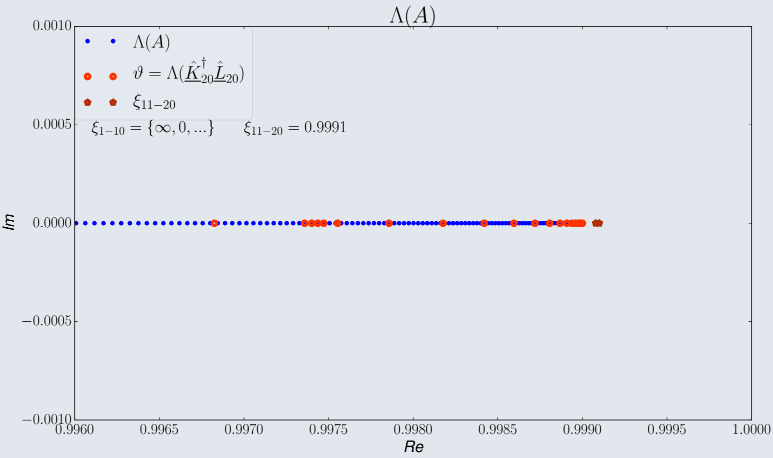


Rational Krylov

Illustrative example

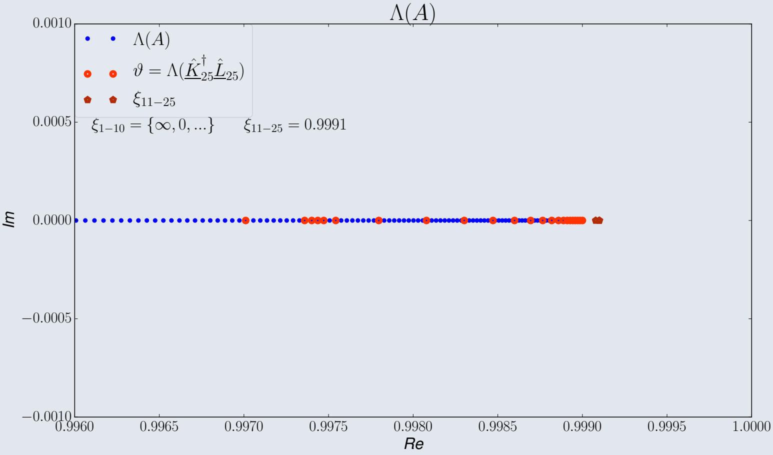


Illustrative example



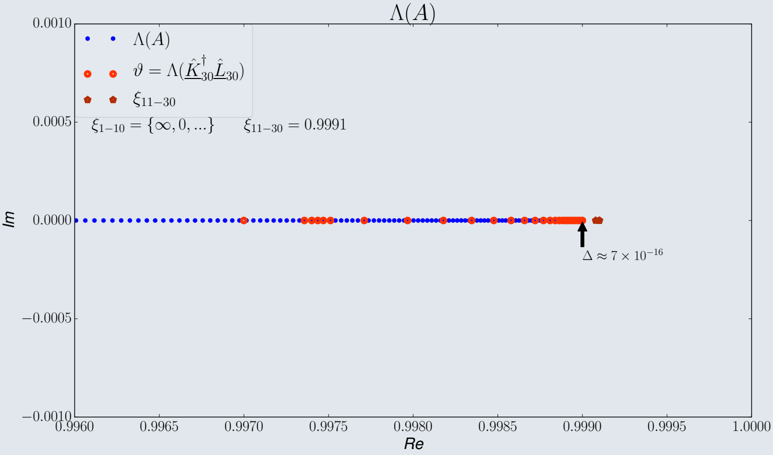
Rational Krylov

Illustrative example



Rational Krylov

Illustrative example



Conclusion

Conclusion

- Filtering (rational) Krylov subspaces is useful to limit subspace dimension when searching for eigenvalues satisfying a property $\mathfrak{P}$
- The filter can be applied implicitly by means of elementary unitary operations (*chasing*) both for polynomial and rational Krylov

Thank you

References

Berljafa, M. and Güttel, S. (2015). Generalized Rational Krylov Decompositions with an Application to Rational Approximation. *SIAM J. Matrix Anal. Appl.*, 36(2):894–916.

Camps, D., Meerbergen, K., and Vandebril, R. (2016). Implicit restart of the extended Krylov method. *Submitted*.

De Samblanx, G., Meerbergen, K., and Bultheel, A. (1997). The implicit application of a rational filter in the RKS method. *BIT Numer. Math.*, 37(4):925–947.

Druskin, V. and Knizhnerman, L. (1998). Extended Krylov Subspaces: Approximation of the Matrix Square Root and Related Functions. *SIAM J. Matrix Anal. Appl.*, 19(3):755–771.

Ruhe, A. (1998). Rational Krylov: A practical algorithm for large sparse nonsymmetric matrix pencils. *SIAM J. Sci. Comput.*, 19(5):1535–1551.

Saad, Y. (1980). Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear Algebra Appl.*, 34:269–295.

Sorensen, D. C. (1992). Implicit Application of Polynomial Filters in a k-Step Arnoldi Method. *SIAM J. Matrix Anal. Appl.*, 13(1):357–385.

Stewart, G. (2001). A Krylov-Schur algorithm for large eigenproblems. *SIAM J. Matrix Anal. Appl.*, 23(3):601–614.