1 Convergence of the Arnoldi method for eigenvalue problems

Recall that, unless it breaks down, k steps of the Arnoldi method generates an orthogonal basis of a Krylov subspace, represented by a matrix $Q = (q_1, \ldots, q_k) \in \mathbb{C}^{n\times k}$ such that $Q^*Q = I$ and

$$\text{span}(q_1, \ldots, q_k) = \mathcal{K}_k(A, b) := \text{span}(b, Ab, \ldots, A^{k-1}b).$$

The eigenvalue approximations (called Ritz values) are subsequently found from the eigenvalues of

$$H = Q^*AQ.$$

The matrix $H \in \mathbb{C}^{k\times k}$ is a Hessenberg matrix and can be generated as a by-product of the Arnoldi method. We call a pair (μ, Qv) a Ritz pair and Qv a Ritz vector, if v and μ safisfy

$$Hv = \mu v.$$

1.1 Bound for subspace-eigenvector angle

As a first indicator of the convergence we will characterize the following quantity

$$\text{error in eigenvector } x_i \sim \|(I - QQ^*)x_i\| \tag{1.1}$$

Recall: $Q \in \mathbb{C}^{n\times k}$ is an orthogonal matrix which means that $Q^*Q = I \in \mathbb{C}^{k\times k}$. However, $I \neq QQ^* \in \mathbb{C}^{n\times n}$.

where

$$Ax_i = \lambda_i x_i.$$

It is very natural to associate the accuracy of the eigenvector with this quantity from a geometric perspective. The indicator in the right-hand side of (1.1) is called (the norm of) the orthogonal complement of the projection of x_i onto the space spanned by Q and it can be interpreted as the sine of the canonical angle between the Krylov subspace and an eigenvector. For the moment, we will only justify this indicator with this geometric reasoning and the following observation:

Lemma 1.1.1 *Suppoe (λ_i, x_i) is an eigenpair A. If the Krylov subspace contains the eigenvector ($x_i \in \mathcal{K}_k(A,b)$), then the indicator vanishes $\|(I - QQ^*)x_i\| = 0$ and there is at least one Ritz value μ such that $\mu = \lambda_i$.*

The Arnoldi method produces an exact approximation if the Krylov subspace contains an eigenvector, or equivalently the indicator is zero.

In words:

- Suppose the Krylov subspace contains the eigenvector ($x_i \in \mathcal{K}_k(A,b)$). Then, there exists a vector $z \in \mathbb{C}^k$ such that $x_i = Qz$. Moreover, this is an eigenvector of H such that the Arnoldi method will generate an exact eigenvalue of A. Moreover, the indicator is $\|(I - QQ^*)x_i\| = \|(I - QQ^*)Qz\| = 0$.
- If, similar to above, $x_i \approx x \in \mathcal{K}_k(A,b)$, we expect the indicator to be small and an eigenvalue of H also to be close λ_i.

The indicator can be bounded as follows, where we assume diagonalizability of the matrix.

Theorem 1.1.2 *Suppose $A \in \mathbb{C}^{n\times n}$ is diagonalizable and let the matrix $X = (x_1, \ldots, x_n) \in \mathbb{C}^{n\times n}$ and diagonal matrix $\Lambda \in \mathbb{C}^{n\times n}$ be the Jordan decomposition such that*

$$A = X\Lambda X^{-1}.$$

Suppose $\alpha_1, \ldots, \alpha_n \in \mathbb{C}\backslash\{0\}$ are such that

$$b = \alpha_1 x_1 + \cdots + \alpha_n x_n \tag{1.2}$$

Recall: The eigenvectors of a diagonalizable matrix form a basis of $\mathbb{C}^n$.

and

$$\varepsilon_i^{(m)} := \min_{\substack{p\in P_{m-1}\\ p(\lambda_i)=1}} \max(|p(\lambda_1)|, \ldots, |p(\lambda_{i-1})|, |p(\lambda_{i+1})|, \ldots, |p(\lambda_n)|)$$

where P_n denotes polynomials of degree n. Suppose the Arnoldi method does not break down when applied to A and started with b. Let $Q \in \mathbb{C}^{n\times m}$ be the orthogonal basis generated after m iterations. Then,

$$\|(I - QQ^*)x_i\| \leq \xi_i \varepsilon_i^{(m)}, \tag{1.3}$$

The indicator can be bounded by a product consisting of two scalar values: $\varepsilon_i^{(m)}$ which only depends on the eigenvalues and iteration number; and ξ_i only depending on the starting vector and eigenvectors.

where

$$\xi_i = \sum_{\substack{j=1\\ j\neq i}}^{n} \frac{|\alpha_j|}{|\alpha_i|}.$$

Proof The proof consists of three steps.

1. Consider any vector $u \in \mathbb{C}^n$. Then

$$\min_{z\in\mathbb{C}^m} \|u - Qz\|_2$$

is a linear least squares problem with a solution given by the normal equations $Q^*u = Q^*Qz$. Hence, $z = Q^*u$. This implies that (for any vector u) we have

$$\min_{z\in\mathbb{C}^m} \|u - Qz\|_2 = \|u - QQ^*u\| = \|(I - QQ^*)u\|$$

2. Although we ultimately want to bound the left-hand side of (1.3), the proof is simplified by considerations of a scaling the left-hand side of (1.3) with α_i as follows:

$$\begin{aligned} \|(I - QQ^*)\alpha_i x_i\| &= \min_{z\in\mathbb{C}^m} \|\alpha_i x_i - Qz\| \\ &= \min_{y\in\mathcal{K}_m(A,b)} \|\alpha_i x_i - y\| \end{aligned}$$

Apply step 1 reversely with $u = \alpha_i x_i$

Now note that the space $\mathcal{K}_m(A,b)$ can be characterized with polynomials. It is easy to verify that $y \in \mathcal{K}_m(A,b)$ is equivalent to the existance of a polynomial $p \in P_{m-1}$ such that $y = p(A)b$. Consequently,

$$\|(I - QQ^*)\alpha_i x_i\| = \min_{p\in P_{m-1}} \|\alpha_i x_i - p(A)b\|.$$

3. The final step consists of inserting the expansion of b in terms of eigenvectors (1.2) and applying appropriate bounds:

$$\begin{aligned} \|(I - QQ^*)\alpha_i x_i\| &= \min_{p\in P_{m-1}} \|\alpha_i x_i - p(A)\sum_{j=1}^{n} \alpha_j x_j\| \\ &= \min_{p\in P_{m-1}} \left\|\alpha_i x_i - \sum_{j=1}^{n} \alpha_j p(\lambda_j) x_j\right\| \\ &\leq \min_{\substack{p\in P_{m-1} \\ p(\lambda_i)=1}} \left\|\alpha_i x_i - \sum_{j=1}^{n} \alpha_j p(\lambda_j) x_j\right\| \\ &= \min_{\substack{p\in P_{m-1} \\ p(\lambda_i)=1}} \|\alpha_i x_i - \alpha_i x_i - \sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_j p(\lambda_j) x_j\| \\ &= \min_{\substack{p\in P_{m-1} \\ p(\lambda_i)=1}} \|\sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_j p(\lambda_j) x_j\| \\ &\leq (\sum_{\substack{j=1 \\ j\neq i}}^{n} |\alpha_j|) \cdot \min_{\substack{p\in P_{m-1} \\ p(\lambda_i)=1}} \max_{j\neq i}(|p(\lambda_j)|) \\ &= (\sum_{\substack{j=1 \\ j\neq i}}^{n} |\alpha_j|) \cdot \varepsilon_i^{(m)} \end{aligned}$$

Since x_i eigenvector, $p(A)x_i = p(\lambda_i)x_i$

For any two sets $S \subset Z$: $\min_{z\in Z} g(z) \leq \min_{z\in S} g(z)$

The conclusion of the theorem is established by dividing the equation by $|\alpha_i|$.

□

Note that $\|b\| = 1$ and $\|x_1\| = \cdots = \|x_n\| = 1$. Hence the coefficients $\alpha_1, \ldots, \alpha_n$ are balanced. In particular they satisfy

$$1 = \|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \leq |\alpha_1| + \cdots + |\alpha_n|.$$

and

$$\xi_i = \frac{1}{|\alpha_i|} \sum_{j=1}^{n} |\alpha_j| - 1 \geq \frac{1}{|\alpha_i|} - 1$$

From this we can easily identify a very good situation and a very bad situation.

- Suppose for all $j \neq i$, $\alpha_j = \delta$ and suppose δ is small. We have that $\xi_i = \frac{(n-1)\delta}{\alpha_i}$. Due to balancing α_i cannot be small. Hence, ξ_i is small, showing fast convergence for this eigenvalue.
- On the other hand, if α_i (the component of the starting vector in the direction of the ith eigenvector) is very small, we have $\xi_i \gg 1$ which implies that the right-hand side of (1.3) is large and we have slow convergence.

This serves as a justification for a more general property.

Rule-of-thumb. Starting vector dependency. The Arnoldi method for eigenvalue problems will "favor" eigenvectors which have large components in the starting vector.

The word "favors" is purposely vague. It should be interpreted as the situation that one observes often in practice, but certainly not always. If we have a particular structure in the matrix or starting vector, we might observe convergence to other eigenvalues.

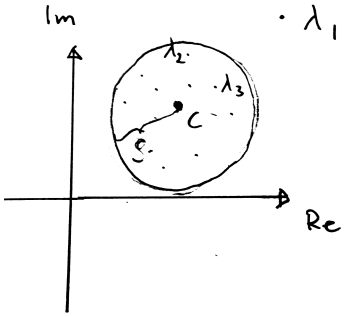
1.1.1 *Bounding* $\varepsilon_i^{(m)}$

In the characterization of the indicator in Theorem 1.1.2 above we introduced the quantity $\varepsilon_i^{(m)}$. This quantity bounds (up to a constant) the error in eigenvector x_i at iteration m. Although $\varepsilon_i^{(m)}$ is defined through a polynomial optimization problem, which is complicated to solve, it is surprisingly easy to use this to obtain bounds providing qualitative understanding of the convergence of the Arnoldi method for eigenvalue problems. We illustrate the power with a specific bound.

Think: $\varepsilon_i^{(m)}$ measures how "difficult" it is to push down a polynomial in points λ_j, for all $j \neq i$ and maintain $p(\lambda_i) = 1$.

Corollary 1.1.3 *Suppose $C(\rho, c) \subset \mathbb{C}$ is a disk centered at $c \in \mathbb{C}$ with radius ρ such that it contains all eigenvalues but λ_1. That is, $\lambda_2, \ldots, \lambda_n \in C(\rho, c)$ and $\lambda_1 \notin C(\rho, c)$. Then,*

$$\varepsilon_1^{(m)} \leq \left(\frac{\rho}{|\lambda_1 - c|} \right)^{m-1}.$$



Proof The proof consists of selecting a particular polynomial in the polynomial optimization problem,

$$\begin{aligned} \varepsilon_1^{(m)} &:= \min_{\substack{p \in P_{m-1} \\ p(\lambda_1)=1}} \max(|p(\lambda_1)|, \ldots, |p(\lambda_{i-1})|, |p(\lambda_{i+1})|, \ldots, |p(\lambda_n)|) \\ &= \max_{j \neq i} |q(\lambda_j)|, \end{aligned}$$

for any $q \in P_{m-1}$ satisfying $q(\lambda_1) = 1$, in particular

$$q(z) = \frac{1}{|\lambda_1 - c|^{m-1}} (z - c)^{m-1}.$$

Hence, from the definition of ρ and c we have that

$$\begin{aligned} \varepsilon_1^{(m)} &\leq \max_{i>1} \frac{|\lambda_i - c|^{m-1}}{|\lambda_1 - c|^{m-1}} \\ &\leq \frac{\rho^{m-1}}{|\lambda_1 - c|^{m-1}}. \end{aligned}$$

□

The result can be intuitively interpreted as follows. If we can construct a small disc that encloses all eigenvalues but one eigenvalue we expect fast (at least linear geometric) convergence for that eigenvalue. This can be achieved for an eigenvalue which is well separated from the rest of the eigenvalues and also in an outer part of the spectrum. We call this "extreme" isolated eigenvalues.

> **Rule-of-thumb. Eigenvalue dependency.** Arnoldi's method for eigenvalue problems favors convergence to "extreme" isolated eigenvalues.

Note the difference between an "extreme" eigenvalue and the eigenvalues which are largest in modulus (absolute value). The Arnoldi method will favor "extreme" whereas the power method will essentially always converge to the eigenvalue largest in modulus.

1.2 *Literature and further reading*

The proof and reasoning above is inspired by [5]. Other convergence bounds involving Schur factorizations, that lead to similar qualitative

understanding can be found in [6], where also complications of the non-generic cases are discussed. There are also further characterizations of convergence and the connection with potential theory [4]. In the above reasoning we characterized the angle between the subspace and the eigenvector. Although this serves as a very accurate prediction of the error in practice, it does not directly give a rigorous bound on the accuracy of Ritz pair. Several approaches to describe the convergence of Ritz values and Ritz vectors have been done in for instance [2, 3]. There is also considerable research on the effect of rounding errors in Krylov methods. Unlike many other numerical methods, the effect of finite arithmetic can improve the performance of the algorithm. See also the recent summary of the convergence of the Arnoldi method for eigenvalue problems [1].

1.3 *Bibliography*

[1] M. Bellalij, Y. Saad, and H. Sadok. Further analysis of the Arnoldi process for eigenvalue problems. *SIAM J. Numer. Anal.*, 48(2):393–407, 2010.

[2] Z. Jia. The convergence of generalized Lanczos methods for large unsymmetric eigenproblems. *SIAM J. Matrix Anal. Appl.*, 16(3):843–862, 1995.

[3] Z. Jia and G. W. Stewart. On the convergence of ritz values, ritz vectors, and refined ritz vectors. Technical report, 1999.

[4] A. B. Kuijlaars. Convergence analysis of Krylov subspace iterations with methods from potential theory. *SIAM Rev.*, 48(1):3–40, 2006.

[5] Y. Saad. *Numerical methods for large eigenvalue problems*. SIAM publications, 2011.

[6] G.W. Stewart. *Matrix Algorithms volume 2: eigensystems*. SIAM publications, 2001.