Controllability of the QR Algorithm on Hessenberg Flags

Jens Jordan, Uwe Helmke
Mathematisches Institut
Universität Würzburg
Am Hubland, D-97074 Würzburg
Germany

Abstract

The shifted QR algorithm can be interpreted as a nonlinear discrete dynamical system on the flag manifold. In the complex case we describe the reachability sets as orbits of a group action and prove non-controllability of the algorithm. In contrast, the algorithm restricted on the subset of Hessenberg flags is generically controllable.

1 Introduction

Spectral shifts are an important tool in numerical linear algebra and serve to speed up the convergence of matrix eigenvalue algorithms. They are e.g. instrumental for the QR algorithm in order to achieve quadratic or cubic convergence rates. Such shift strategies are usually found in a heuristic way and a systematic analysis and design of shift policies is essentially an open research problem in numerical analysis; see [7]. From a control theoretic point of view, matrix eigenvalue algorithms can be viewed as dynamical systems with spectral shifts acting as control variables. In [5, 6], reachability properties of the shifted inverse power iteration on projective space are investigated. Here we extend this earlier approach to a control theoretic analysis of the shifted QR algorithm.

The starting point for our analysis is the well known basic fact that the QR algorithm for computation of eigenspaces can be interpreted as the inverse power iteration on the flag manifold, i.e. as a nonlinear dynamical system on the flag manifold; see e.g. [1]. In this paper we focus on the controllability properties of the shifted inverse power iteration on complex flag manifolds. Reachable sets are shown to be orbits of an abelian group that acts on the flag manifold. This immediately leads to a sharp upper bound on the dimensions of the reachable sets as well as to the conclusion that the shifted QR algorithm for $n \times n$ matrices is never controllable, unless $n = 2$. Moreover, a coarser partition of the flag manifold by unions of reachable sets of a fixed type is described, thus making contact with earlier work by Gelfand et. al. [3] on Grassmann simplices.

In practical implementations of the algorithm an initial matrix $\mathbf{A}$ is first reduced to Hessenberg form. Since the QR algorithm preserves the Hessenberg structure it makes sense to restrict the algorithm to matrices of that type. In our setting this is equivalent to consider the restriction of the shifted inverse power iteration on the set of so-called Hessenberg flags. Hessenberg flags $\mathcal{V} = (V_1, \dots, V_{n-1})$ are increasing sequences of linear subspaces $V_1 \subset \cdots \subset V_{n-1}$ with the property that $\mathbf{A}V_k \subset V_{k+1}$, $k = 1, \dots, n-1$. The set of Hessenberg flags is a compact algebraic subvariety of the flag manifold, called the Hessenberg variety. Its geometry has been studied by de Mari and Shayman [2]. As the Hessenberg variety is invariant under iterations of the shifted inverse iteration, it is in particular a union of reachable sets. We prove that there is one reachable set that is dense in the set of complex Hessenberg flags. This implies that the QR algorithm with complex eigenvalue shifts is generically controllable on the set of Hessenberg matrices. A similar characterization of reachable sets for the shifted QR algorithm on real symmetric isospectral tridiagonal matrices has been given in unpublished work by G. Gladwell and presented in a lecture at the 4th SIAM Conference on Linear Algebra in Signals, Systems and Control, Boston 2001. Gladwell's approach is quite different to ours and is based on the preservation of total positivity of tridiagonal matrices under the QR algorithm; see [4]. We believe that our approach has certain advantages that simplifies the analysis.

2 Reachable sets as orbits of a group action

We begin, by interpreting the QR algorithm as an inverse iteration on the flag manifold. Recall, that a *complex flag* $\mathcal{V}$ is a sequence of complex linear subspaces $V_1 \subset V_2 \subset \cdots \subset V_{n-1}$ with $\dim_{\mathbb{C}} V_k = k$ for all $k = 1, \dots, n-1$. Let $\mathrm{Flag}(\mathbb{C}^n)$ denote the set of all such flags. It is a smooth, compact and connected manifold of complex dimension $\frac{1}{2}n(n-1)$.

The general linear group $\mathrm{GL}_n(\mathbb{C})$ acts on $\mathrm{Flag}(\mathbb{C}^n)$ via

$$\mathcal{V} \to g \cdot \mathcal{V} := (gV_1, gV_2, \dots, gV_{n-1}) \in \mathrm{Flag}(\mathbb{C}^n), \qquad g \in \mathrm{GL}_n(\mathbb{C}).$$

Let $\mathbf{A}$ denote a complex $n \times n$ matrix with spectrum $\sigma(\mathbf{A})$. Following Ammar and Martin [1] we interpret the shifted QR algorithm, acting on linear subspaces, as the nonlinear discrete time control system on $\mathrm{Flag}(\mathbb{C}^n)$, given as

$$\mathcal{V}_{k+1} = (\mathbf{A} - u_k\mathbf{I})^{-1}\mathcal{V}_k, \qquad u_k \notin \sigma(\mathbf{A}) \tag{2.1}$$

Starting from any point $\mathcal{V}_0$ of the flag manifold, the set of flags that can be obtained by a finite number of iteration steps of (2.1), is the *reachable set*

$$\mathcal{R}_A(\mathcal{V}_0) := \{\mathcal{V} \in \mathrm{Flag}(\mathbb{C}^n) \,|\, \mathcal{V} = \Pi_{k=0}^{N}\big(\mathbf{A} - u_k\mathbf{I}\big)^{-1}\mathcal{V}_0 \,|\, N \in \mathbb{N}; u_k \in \mathbb{C} \setminus \sigma(\mathbf{A})\}.$$

The next result shows that reachable sets of (2.1) are orbits of an abelian group action on the flag manifold. Recall that a matrix $\mathbf{A}$ is called *cyclic* if its minimal polynomial coincides with the characteristic polynomial.

Lemma 2.1. *(a) Given $\mathbf{A} \in \mathbb{C}^{n\times n}$, the set*

$$\Gamma_A := \{\Pi_{k=0}^{N}\big(\mathbf{A} - u_k\mathbf{I}\big)^{-1} \mid N \in \mathbb{N}; u_k \in \mathbb{C} \setminus \sigma(\mathbf{A})\}$$

is a abelian subgroup of $\mathrm{GL}_n(\mathbb{C})$. If $\mathbf{A}$ is cyclic, then Γ_A is the centre of $\mathbf{A}$ in $\mathrm{GL}_n(\mathbb{C})$. If $\mathbf{A}$ is diagonalizable, then

$$\Gamma_A = \{\Pi_{k=0}^{n-1}\big(\mathbf{A} - u_k\mathbf{I}\big) \mid u_k \in \mathbb{C} \setminus \sigma(\mathbf{A})\}.$$

(b) For every $\mathcal{V} \in \mathrm{Flag}(\mathbb{C}^n)$, the reachable sets coincide with the orbits of Γ_A:

$$\mathcal{R}_A(\mathcal{V}) = \Gamma_A \cdot \mathcal{V}$$

From the above lemma, if $\mathbf{A}$ is diagonalizable, then every reachable set is the image of a connected set under the continuous map $F_{\mathcal{V}} : (\mathbb{C} \setminus \sigma(\mathbf{A}))^n \to \mathrm{Flag}(\mathbb{C}^n)$

$$F_{\mathcal{V}}(u_0, \ldots, u_{n-1}) = \Pi_{k=0}^{n-1}\big(\mathbf{A} - u_k\mathbf{I}\big)\mathcal{V}.$$

Therefore the image is connected, too. This shows that reachable sets are always connected. Together with the description of reachable sets as orbits of a Lie group action, the following result is obtained. Note that the given dimension bound is sharp.

Theorem 2.1. *If $\mathbf{A}$ is diagonalizable, the reachable sets $\mathcal{R}_A(\mathcal{V})$ are connected complex submanifolds of $\mathrm{Flag}(\mathbb{C}^n)$ of dimension at most $n-1$.*

In particular, every reachable set has dimension at most $n-1$. Since $\mathrm{Flag}(\mathbb{C}^n)$ has dimension $\frac{1}{2}n(n-1)$ we conclude

Corollary 2.1. *Let $n > 2$. The shifted inverse iteration (2.1) on the flag manifold is not controllable. In particular, the shifted QR algorithm on isospectral complex $n \times n$ matrices is not controllable.*

3 Classification of reachable sets

The results of the previous section show that the partition of the flag manifold by reachable sets is a singular foliation by low dimensional submanifolds. Although the geometry of such leaves can be easily described, the partition is to fine to be useful for classification purposes. We therefore introduce a coarser partition of the flag manifold by unions of reachable sets of a fixed type.

Definition 3.1. *For $\mathbf{A} \in \mathbb{C}^{n\times n}$ let $Inv_{\mathbf{A}}$ denote the set of proper $\mathbf{A}$-invariant subspaces $W \subset \mathbb{C}^n$. Two flags $\mathcal{U}, \mathcal{V} \in \mathrm{Flag}(\mathbb{C}^n)$ are called equivalent, $\mathcal{U} \simeq \mathcal{V}$, if for all $W \in Inv_{\mathbf{A}}$:*

$$\dim(U_j \cap W) = \dim(V_j \cap W), \quad j = 1, \ldots, n-1.$$

The set of all flags, that are equivalent to a given flag $\mathcal{U}$ is denoted by $[\mathcal{U}]$.

Theorem 3.1. *$\simeq$ is an equivalence relation on $\mathrm{Flag}(\mathbb{C}^n)$. The equivalence classes are unions of reachable sets. If* $\mathbf{A}$ *is cyclic, there exist only finitely many equivalence classes.*

Note that each equivalence class $[\mathcal{U}]$ has the following equivalent description as an intersection of Schubert cells. For simplicity we assume that $\mathbf{A}$ has distinct eigenvalues with eigenvectors $e_1, \dots, e_n$. For any permutation $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$ let $\mathcal{E}^\pi = (E_1^\pi, \dots, E_{n-1}^\pi)$ denote the flag defined by $E_i^\pi = span\{e_{\pi(1)}, \dots, e_{\pi(i)}\}$. Then

$$S_{\mathcal{U}}(\mathcal{E}^\pi) := \{\mathcal{V} \in \mathrm{Flag}(\mathbb{C}^n) \mid \dim(V_i \cap E_j^\pi) = \dim(U_i \cap E_j^\pi)\ \forall i, j\}$$

is a Schubert cell in the flag manifold and

$$[\mathcal{U}] = \bigcap_\pi S_{\mathcal{U}}(\mathcal{E}^\pi)$$

is the intersection of these $n!$ Schubert cells. Such objects have been first studied by Gelfand et. al. [3] and are called Grassmann simplices. The filtration by Grassmann simplices is coarser then that by reachable sets.

From $\max\{0, \dim(V) + \dim(W) - n\} \leq \dim(V \cap W)$ we expect that the largest Grassmann simplex is

$$\begin{aligned} \mathcal{M}_{\mathbf{A}} \ := \ \big\{&\mathcal{V} \in \mathrm{Flag}(\mathbb{C}^n) \mid \dim(V_k \cap W) = \max\{0, \dim(V_k) + \dim(W) - n\} \\ &\text{for all } W \in Inv_{\mathbf{A}} \text{ and } k = 1, \dots, n-1\big\}. \end{aligned}$$

For generic matrices $\mathbf{A}$ this is indeed true:

Theorem 3.2. *If $\mathbf{A} \in \mathrm{GL}_n(\mathbb{C})$ has n distinct eigenvalues, then $\mathcal{M}_A$ is open and dense in $\mathrm{Flag}(\mathbb{C}^n)$.*

Example 3.1. *Let $\mathbf{A} \in \mathrm{GL}_3(\mathbb{C})$ have distinct eigenvalues. If e_1, e_2, e_3 is a basis of eigenvectors of* $\mathbf{A}$ *then*

$$Inv_{\mathbf{A}} = \{\langle e_1 \rangle, \langle e_2 \rangle, \langle e_3 \rangle, \langle e_1, e_2 \rangle, \langle e_1, e_3 \rangle, \langle e_2, e_3 \rangle\}.$$

Each equivalence class of a flag $\mathcal{U} \in \mathrm{Flag}(\mathbb{C}^3)$ ist characterized by 12 numbers $\dim(U_i \cap W) \in \{0, 1, 2\}$. The class $\mathcal{M}_A$ is the set of flags $\mathcal{U} = (U_1, U_2)$ with $\dim(U_1 \cap W) = 0$ and $\dim(U_2 \cap W) = \dim(W) - 1$.

The following conjecture describes the adherence order on Grassmann simplices.

Conjecture 3.1. *Let $\mathbf{A} \in \mathrm{GL}_n(\mathbb{C})$ have n distinct eigenvalues and $\mathcal{U}, \mathcal{V} \in \mathrm{Flag}(\mathbb{C}^n)$, then*

$$[\mathcal{U}] \subseteq \overline{[\mathcal{V}]} \quad \Leftrightarrow \quad \textit{for all } \ W \in Inv_{\mathbf{A}} : \ \dim(U_j \cap W) \geq \dim(V_j \cap W),\ j = 1, \dots, n-1.$$

4 Reachable sets of the Hessenberg variety

In numerical computations one transforms a matrix $\mathbf{A}$ first into Hessenberg form and then applies the QR algorithm to this condensed form. Since the QR algorithm preserves the Hessenberg structure it restricts to a control system on the set of Hessenberg flags; see [1] and [2].

For a given matrix $\mathbf{A}$, the Hessenberg variety is defined as the set

$$\mathrm{Hess}_A(\mathbb{C}^n) := \{\mathcal{V} \in \mathrm{Flag}(\mathbb{C}^n) \,|\, \mathbf{A}V_k \subset V_{k+1},\, k = 1, \ldots, n-1\}.$$

De Mari and Shayman [2] have shown that $\mathrm{Hess}_A(\mathbb{C}^n)$ is a compact connected submanifold of $\mathrm{Flag}(\mathbb{C}^n)$, provided $\mathbf{A}$ has n distinct eigenvalues. Moreover, the dimension in this case is $n-1$. On the other hand, the Hessenberg variety is invariant under iterations of the control system (2.1) and therefore it is a union of reachable sets. Since these reachable sets have at most dimension $n-1$, one expects that there is a dense reachable set in $\mathrm{Hess}_A(\mathbb{C}^n)$. This is indeed true and shown below.

The following result characterizes the intersection of $\mathcal{M}_A$ with the Hessenberg variety.

Lemma 4.1. *Let $\mathbf{A}$ have n distinct eigenvalues and $\mathcal{V} \in \mathrm{Hess}_A(\mathbb{C}^n)$. Then $\mathcal{V} \in \mathcal{M}_A$ if and only if $\mathbf{A}V_k \neq V_k$ for all $k = 1, \ldots, n-1$. In particular, $\mathcal{M}_A \cap \mathrm{Hess}_A(\mathbb{C}^n)$ is an open and dense subset of* $\mathrm{Hess}_A(\mathbb{C}^n)$.

Note that the set $\mathrm{Hess}_A(\mathbb{C}^n) \cap \mathcal{M}_A$ is invariant under the algorithm and therefore must be a union of reachable sets. It is actually equal to a reachable set.

Theorem 4.1. *Let $\mathbf{A} \in \mathrm{GL}_n(\mathbb{C})$ have distinct eigenvalues. For every $\mathcal{V} \in \mathrm{Hess}_A(\mathbb{C}^n) \cap \mathcal{M}_A$ then $\mathcal{R}_A(\mathcal{V}) = \mathrm{Hess}_A(\mathbb{C}^n) \cap \mathcal{M}_A$.*

This implies the main result of this paper.

Theorem 4.2. *Let $\mathbf{A} \in \mathrm{GL}_n(\mathbb{C})$ have n distinct eigenvalues. For every $\mathcal{V} \in \mathrm{Hess}_A(\mathbb{C}^n) \cap \mathcal{M}_A$*

$$\overline{\mathcal{R}_A(\mathcal{V})} = \mathrm{Hess}_A(\mathbb{C}^n).$$

In particular, the shifted inverse iteration (2.1), restricted to the Hessenberg variety, is controllable.

It can be shown, using the above result, that the shifted QR algorithm on complex isospectral Hessenberg matrices is controllable. In the real case the situation is a bit more complicated and only partially understood. If $\mathbf{A}$ is a real symmetric matrix with distinct eigenvalues, then (2.1) with real shifts is controllable on the real Hessenberg variety. This implies controllability of the shifted QR algorithm on the set of symmetric isospectral tridiagonal matrices, with positive off diagonal entries. There are several open research problems in this area. One is the characterization of the adherence order for the reachable sets, i.e. a characterization of the reachable sets contained in the closure of a given one. The extension of our results to FG algorithms is another challenge, as is the extension of the algorithm to other classical groups or general semisimple Lie groups.

5 Conclusions

The QR algorithm with eigenvalue shifts on for eigenspace computations can be equivalently reformulated as the shifted inverse iteration on the flag manifold. It is never controllable, except for 2×2 matrices. For arbitrary complex matrices with distinct eigenvalues we show that the shifted inverse iteration on the Hessenberg variety is controllable. This implies controllability of the shifted QR algorithm on complex isospectral Hessenberg matrices.

References

[1] G. Ammar, C. Martin, The Geometry of Matrix Eigenvalue Methods. *Acta Applicandae Mathematicae* **5** (1986), 239-278.

[2] F. De Mari, M. Shayman, Generalized Eulerian numbers and the topology of the Hessenberg variety of a matrix. *Acta Applicandae Mathematicae* **12** (1988), 213-235.

[3] I.M. Gelfand, R.M. Goresky, R.D. MacPherson, V.V. Serganova, Combinatorial geometries, convex polyhedra and Schubert cells. *Advances in Mathematics* **63** (1987), 301-316.

[4] G. M. L. Gladwell, Total Positivity and the QR Algorithm. *Linear Algebra and its Applications* **271** (1998), 257-272.

[5] U. Helmke, P.A. Fuhrmann, Controllability of matrix eigenvalue algorithms: the inverse power method. *Systems and Control Letters* **41** (2000), 57-66.

[6] U. Helmke, F. Wirth, On controllability of the real shifted inverse power iteration. *Systems and Control Letters* **43** (2001), 9-23.

[7] P. Van Dooren, R. Sepulchre, Shift policies in QR-like algorithms and feedback control of self-similar flows. *Open Problems in Mathematical Systems and Control Theory* (V.D. Blondel, E.D. Sontag, M. Vidyasagar and J.C. Willems, Eds.), Communications and Control Engineering Series. Springer Publ., London, 1999, 245-249.