Well-posedness and regularity properties of the Grassmann-Rayleigh quotient iteration

P.-A. Absil* U. Helmke† K. Hüper‡

Technical Report FSU-CSIT-04-26
December 9, 2004

Abstract. A generalization of the Rayleigh quotient iteration has recently been proposed on the Grassmann manifold. This iteration has been shown to converge locally cubically to the invariant subspaces of symmetric matrices. The present paper studies global properties of the iteration mapping. Results are obtained e.g. concerning fixed points, smoothness, and singularities of the iteration mapping.

Key words. Grassmann-Rayleigh quotient iteration, Block-Rayleigh quotient iteration, Grassmann manifold, singularities, continuous extension, fixed points.

AMS subject classification. 65F15.

1 Introduction

The Rayleigh quotient iteration (RQI) is a well-known method for computing an eigenvector of a symmetric matrix. It consists in a shifted inverse iteration that maps $y \in \mathbb{R}^n \backslash \{0\}$ to y_+ defined by

$$(A - \rho(y)I)z = y, \; y_+ = \mu z \tag{1}$$

where μ is a normalization factor and $\rho(y) := (y^T Ay)/(y^T y)$ is the Rayleigh quotient of A evaluated at y.

The following $p \times p$ block-shift generalization of the RQI method has recently been pro-

*School of Computational Science and Information Technology, Florida State University, Tallahassee FL 32306-4120, USA (www.csit.fsu.edu/∼absil). This author's work was supported partially by the School of Computational Science of Florida State University under a postdoctoral fellowship. Part of this work was done while the author was a Research Fellow with the Belgian National Fund for Scientific Research (Aspirant du F.N.R.S.) at the University of Liège. This work was partially performed while the author was a guest in the Department of Mathematics, University of Würzburg, under a grant from the European Nonlinear Control Network. The hospitality of the members of the department is gratefully acknowledged.

†Department of Mathematics, University of Würzburg, 97074 Würzburg, Germany.

‡National ICT Australia, Canberra Research Laboratory, SEACS Program, Locked Bag 8001, Canberra ACT 2601, Australia (knut.hueper@nicta.com.au). National ICT Australia is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence Program.

posed in [AMSV02][1]. It maps the full-rank $n \times p$ matrix Y to Y_+ defined by

$$AZ - Z(Y^TY)^{-1}Y^TAY = Y, \tag{2a}$$

$$Y_+ = ZM \tag{2b}$$

where Z and Y_+ are $n \times p$ matrices and M is any invertible $p \times p$ matrix used for normalization purposes. This iteration induces an iteration on the set of p-dimensional subspaces of $\mathbb{R}^n$. That is, the column space of Y_+, denoted by $\mathrm{span}(Y_+)$, only depends on $\mathrm{span}(Y)$. The set of p-dimensional subspaces of $\mathbb{R}^n$ is termed the *Grassmann manifold*, denoted here by $\mathrm{Grass}(p, n)$, hence the name *Grassmann-RQI* (GRQI) for the subspace iteration induced by (2).

The RQI method (which has been known for more than half a century, see [Cra51]) was initially used as a way of improving an initial estimate of an eigenvector. In a series of papers dedicated to the iteration, Ostrowski [Ost59] showed that the RQI method converges cubically. This result is local in nature: under generic conditions, if the iteration is started sufficiently close to an eigenvector, then it converges to the eigenvector, and does so with cubic rate; see e.g. the proofs in [Par74, Par80] or the calculus-based proof in [Hüp03].

This local convergence result extends to the GRQI method. Locally around the isolated[2] p-dimensional invariant subspaces of A, the singularities in (2) can be removed by continuous extension and the iteration converges with cubic rate; see [AMSV02, Abs03].

It is natural to ask how the iteration performs if it is started from an arbitrary initial condition. In the scalar case ($p = 1$), the global behaviour of the iteration is well understood. A proof due to Parlett and Kahan (see [Par74, Par80]) shows that the sequence of iterates converges either to an eigenvector or to a periodic orbit of period two. Using dynamical systems techniques, Batterson and Smillie [BS89] have shown that it converges to an eigenvector for almost all initial condition.

In contrast to the scalar Rayleigh quotient iteration, the block-shift generalization (2) involves additional complications due to a loss of well-definedness of the iteration. In the scalar case, inverse Rayleigh iteration is only well defined as long as the Rayleigh quotient differs from an eigenvalue of A. This creates a possible singularity for the algorithm which has to be analyzed. In the matrix case, a new feature enters through the possibility of a rank drop, even at a regular point. While in the scalar case a rather complete analysis of the dynamics around such singular points is possible, the matrix case poses further challenges towards the development of a complete phase portrait analysis. The currently available results on the global convergence of the Grassmann-RQI are mainly based on numerical experiments. In [ASVM04], it was shown how the basins of attraction of the invariant subspaces depend on the distribution of the eigenvalues of A, and numerical convergence to an invariant subspace was always observed for randomly chosen initial condition.

The objective of the paper is to address these issues by initiating an analytical study of the singularities of the GRQI algorithm. In Section 2, the two kinds of singularities that may

[1] After [AMSV02] was published the authors became aware that the iteration had been previously considered in the PhD thesis [Smi97] and the manuscript [ES99]. Smit [Smi97] gives a proof of cubic convergence that significantly differs from the one in [AMSV02]. The former relies on the cubic convergence of the classical RQI while the latter uses a well-chosen coordinate system on the Grassmann manifold.

[2] Since A is symmetric, its isolated invariant subspaces are those for which the eigenvalues of A restricted to the invariant subspace are disjoint from the eigenvalues of A restricted to the orthogonal complement of the invariant subspace; see e.g. [RR02]. These invariant subspaces are called *simple* in [SS90].

appear in the GRQI equation (2) are presented. An extended GRQI mapping, defined everywhere on the Grassmann manifold, is proposed in Section 3. The second kind of singularity, i.e. rank deficiency in Z, is studied in Section 4. Section 5 is concerned with smoothness of the iteration mapping. Section 6 deals with fixed points, periodic orbits and other global convergence issues. Finally, Section 7 summarizes the main results and presents open questions.

2 Singularities

It is assumed that Y in (2) is full-rank, i.e. Y belongs to the *noncompact Stiefel manifold*

$$\mathrm{ST}(p,n) := \{Y \in \mathbb{R}^{n\times p} : \det(Y^TY) \neq 0\}.$$

We also assume throughout that the computations are done in exact arithmetic.

Two kinds of singularities may occur in the iteration defined by (2). A first kind of singularity occurs when the equation (2a) does not admit one and only one solution. This happens if and only if A and the block-shift $R_A(Y) := (Y^TY)^{-1}Y^TAY$ have at least one eigenvalue in common. Note that if a singularity of the first kind occurs at Y, then it also occurs everywhere on the set $\mathrm{span}^{-1}(\mathrm{span}(Y)) = \{YM : \det(M) \neq 0\}$ containing all the $n \times p$ matrices that have the same column space as Y. Therefore, we say that singularity of the first kind is a Grassmannian property and we define

$$\psi := \{\mathrm{span}(Y) : Y \in \mathrm{ST}(p,n) \text{ and } \mathrm{spec}(A) \cap \mathrm{spec}((Y^TY)^{-1}Y^TAY) \neq \emptyset\}, \tag{3}$$

where $\mathrm{spec}(A)$ denotes the set of the eigenvalues of A. When $\mathrm{span}(Y) \notin \psi$, we denote by $Z(Y)$ the unique solution of (2a).

A second kind of singularity occurs if Z is rank-deficient because the next iterate would not belong to $\mathrm{Grass}(p,n)$, but to some $\mathrm{Grass}(\alpha,n)$ with $\alpha < p$. Since $Z(YM) = Z(Y)M$, it follows that rank deficiency is also a Grassmannian property. In Section 4, we will prove for $p \in \{1,2\}$ that rank-deficiency does not occur outside singularities of the first kind. Moreover, we will give numerical evidence that this property also holds for $p \geq 3$. We thus claim that

$$\mathcal{RD} := \{\mathrm{span}(Y) \in \mathrm{Grass}(p,n)\backslash\psi : Z(Y) \text{ is rank-deficient}\} \tag{4}$$

is empty, or in other words:

Conjecture 2.1 ($\mathcal{RD} = \emptyset$) *Let A be a symmetric $n \times n$ matrix. Let X be an orthonormal $n \times p$ matrix such that the eigenvalues of X^TAX are disjoint from the eigenvalues of A. Let Z be the solution of*

$$AZ - ZX^TAX = X, \tag{5}$$

which exists and is unique. Then Z is full-rank.

We point out that (5) is a special form of the general Sylvester equation $AZ - ZB = C$, whose properties have been analyzed in great detail in the literature [Hea77, dSB81, Wim88]. In particular, sufficient conditions have been given on (A, B, C) for the solution Z to have full rank. However, these sufficient conditions are not satisfied in (5)—for example, the condition in [Hea77] that C has rank one obviously does not apply here since X in (5) is orthonormal.

It is interesting to note that in general, the solution of a Sylvester equation $AZ - ZB = C$ can be rank deficient, even when A, B and C are full-rank. In particular, if the classical RQI mapping is applied in parallel to linearly independent vectors, then the outputs may not be linearly independent, as one of us shows in [Hüp02, Hüp03] with the following example: Let

$$A = \begin{bmatrix} 1 & \sqrt{2} & 0 \\ \sqrt{2} & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

with eigenvalues $\lambda_1 = -1$ and $\lambda_{2,3} = \frac{3}{2} \pm \sqrt{\frac{5}{4}}$, let C be the 3-by-3 identity matrix and let $B = \mathrm{diag}(1,1,0)$ be the diagonal component of $C^T AC$; then $AZ - ZB = C$ has for unique solution

$$Z = \begin{bmatrix} -1/2 & \sqrt{2}/2 & -\sqrt{2} \\ \sqrt{2}/2 & 0 & 1 \\ \sqrt{2}/2 & 0 & 1 \end{bmatrix},$$

which clearly has only rank 2.

Before considering this question of rank-deficiency in more details in Section 4, we now define an extended GRQI mapping that is also defined on ψ.

3 Extended GRQI mapping

The classical RQI mapping (1) is well defined on $\mathrm{Grass}(1,n) \setminus \psi$, i.e. the shift must not be an eigenvalue of A. Extensions of the RQI mapping that are defined on the whole projective space $\mathrm{Grass}(1,n)$ have been proposed in [BS89, PS95]. These extensions coincide with the RQI mapping or its continuous extension everywhere they exist. In this section, we draw inspiration from the $p = 1$ case to derive an extended GRQI mapping. Roughly speaking, when an eigenvalue of the block shift $R_A(Y) := (Y^TY)^{-1}Y^TAY$ is equal to an eigenvalue of A, the extended mapping performs a projection onto the corresponding invariant subspace of A.

In order to formally define the extended GRQI mapping, we first introduce some notation. Without loss of generality since $A = A^T$, we work in a coordinate system in which $A = \mathrm{diag}(\lambda_1 I_{\mu_1}, \ldots, \lambda_r I_{\mu_r})$, $\lambda_1 < \ldots < \lambda_r$, with λ_i occuring with multiplicity μ_i, $\mu_1 + \ldots + \mu_r = n$. Then the canonical vectors e_i, $i = 1, \ldots, n$, form an orthonormal basis of eigenvectors of A. Let $\mathcal{E}_i$ denote the full μ_i-dimensional invariant subspace of A relative to λ_i, $i = 1, \ldots, r$. Let $\mathrm{St}(p,n)$ denote the compact Stiefel manifold, i.e. the set of orthonormal $n \times p$ matrices. Let $\mathcal{S}$ be a p-dimensional subspace of $\mathbb{R}^n$, i.e. an element of $\mathrm{Grass}(p,n)$. Denote by $P_{\mathcal{S}}$ the orthogonal projection onto $\mathcal{S}$ and let

$$A_{\mathcal{S}} : \mathcal{S} \to \mathcal{S} : y \mapsto P_{\mathcal{S}} A y$$

be the compression of A to $\mathcal{S}$. Note that $A_{\mathcal{S}}$ is a self-adjoint operator. Let $\rho_1 < \ldots < \rho_s$ be the eigenvalues of $A_{\mathcal{S}}$ with multiplicities $\nu_1, \ldots, \nu_s$, $\nu_1 + \ldots + \nu_s = p$ and let $\mathcal{S}_1, \ldots, \mathcal{S}_s$ be the corresponding invariant subspaces, of dimension $\nu_1, \ldots, \nu_s$, respectively. The ρ's are called the *Ritz values* of $(A, \mathcal{S})$ and the $\mathcal{S}_i$'s are the corresponding *Ritz spaces.* If $X^TX = I$ and $\mathrm{span}(X) = \mathcal{S}$, then the Ritz values of $(A, \mathcal{S})$ are the eigenvalues of X^TAX, with the same multiplicity.

Using these definitions and notation we define an *extended GRQI mapping* as follows:

Definition 3.1 (F_{GRQI} – extended GRQI mapping) *Given $A = A^T \in \mathbb{R}^{n\times n}$, F_{GRQI} maps $\mathcal{S} \in \mathrm{Grass}(p,n)$ to $\mathcal{S}^+ \in \cup_{\alpha=0}^p \mathrm{Grass}(\alpha,n)$ as follows:*
a. Compute the Ritz values $\rho_1 < \ldots < \rho_s$ and the corresponding Ritz spaces $\mathcal{S}_1,\ldots,\mathcal{S}_s$ of $(A,\mathcal{S})$.
b. For $i = 1,\ldots,s$:

$$\mathcal{S}_i^+ := \begin{cases} (A-\rho_i I)^{-1}\mathcal{S}_i & \text{if } \rho_i \notin \mathrm{spec}(A) \\ P_{\mathcal{E}_j}\mathcal{S}_i & \text{if } \rho_i = \lambda_j \end{cases} \tag{6}$$

where $P_{\mathcal{E}_j}$ is the orthogonal projection onto the kernel of $(A-\lambda_j I)$.
c. $\mathcal{S}^+ := \oplus_{i=1}^s \mathcal{S}_i^+$.

We now formulate this mapping as a matrix algorithm.

Algorithm 3.2 (extended GRQI) *Data: $A = A^T \in \mathbb{R}^{n\times n}$.*
Input: $\mathcal{S} \in \mathrm{Grass}(p,n)$.
Output: $\mathcal{S}^+ \in \cup_{\alpha=0}^p \mathrm{Grass}(\alpha,n)$.
a. Pick $\tilde{X} \in \mathrm{St}(p,n)$ such that $\mathrm{span}(\tilde{X}) = \mathcal{S}$. Find $V \in O(p)$ such that $V^T\tilde{X}^TA\tilde{X}V = \mathrm{diag}(\rho_1,\ldots,\rho_1,\ldots,\rho_s,\ldots,\rho_s)$ with $\rho_1 < \ldots < \rho_s$ appearing with multiplicity $\nu_1,\ldots,\nu_s$ respectively. Let $X := \tilde{X}V$. Note that V is defined up to post-multiplication by any element of $O(\nu_1)\times\ldots\times O(\nu_n)$ and so is X. Decompose X as $(X_1|\ldots|X_s)$ where X_i has ν_i columns, and let x_j denote the jth column of X. The x_i's are called Ritz vectors *and their corresponding Ritz values are $x_i^TAx_i$ in $\mathrm{spec}(A,\mathcal{S})$.*
b. For $i = 1,\ldots,s$:

$$Z_i := \begin{cases} (A-\rho_i I)^{-1}X_i & \text{if } \rho_i \notin \mathrm{spec}(A) \\ P_{\mathcal{E}_j}X_i & \text{if } \rho_i = \lambda_j \end{cases} \tag{7}$$

where $P_{\mathcal{E}_j}$ is the orthogonal projection onto the kernel of $(A-\lambda_j I)$.
c. $\mathcal{S}^+ := \mathrm{span}(Z)$ where $Z := (Z_1|\ldots|Z_s)$.

Comparing Definition 3.1 and Algorithm 3.2, one has $\mathcal{S}_i = \mathrm{span}(X_i)$ and $\mathcal{S}_i^+ = \mathrm{span}(Z_i)$.

The definition (7) chosen for Z_i is justified by the following result, which shows that the function Z_i of the arguments (ρ_i, X_i) defined in (7) is smooth up to a scaling factor (which does not modify the span of Z_i). This result will also be useful when we study the smoothness of F_{GRQI} in Theorem 5.1.

Lemma 3.3 *Given $A = A^T \in \mathbb{R}^{n\times n}$ with eigenvalues $\lambda_1 < \ldots < \lambda_r$ and λ_i occuring with multiplicity μ_i, $i = 1,\ldots,r$, let*

$$z_{\rho,x} := \begin{cases} (A-\rho I)^{-1}x & \text{if } \rho \notin \mathrm{spec}(A), \\ P_{\mathcal{E}_j}x & \text{if } \rho = \lambda_j \end{cases} \tag{8}$$

where $P_{\mathcal{E}_j}$ is the orthogonal projection onto the kernel of $(A-\lambda_j I)$, and let

$$z'_{\rho,x} := \alpha_\rho z_{\rho,x} \tag{9}$$

where α_ρ is a scalar defined by

$$\alpha_\rho := \begin{cases} \prod_{1\le i\le r}(\lambda_i - \rho) & \text{if } \rho \notin \mathrm{spec}(A), \\ \prod_{\substack{1\le i\le r \\ i\ne j}}(\lambda_i - \rho) & \text{if } \rho = \lambda_j. \end{cases} \tag{10}$$

Then $z'_{\rho,x}$ is a smooth function of its arguments.

Proof. (Lemma 3.3) Let $A = VDV^T$ be an eigenvalue decomposition of A, with D containing the eigenvalues of A in increasing order. It comes from (8), (9), (10) that for all ρ and all x,

$$z'_{\rho,X} = V\mathrm{diag}(\underbrace{\prod_{\substack{1\le i\le r\\ i\ne 1}}(\lambda_i-\rho),\ldots,\prod_{\substack{1\le i\le r\\ i\ne 1}}(\lambda_i-\rho)}_{\mu^1\text{elements}},\ldots,\underbrace{\prod_{\substack{1\le i\le r\\ i\ne r}}(\lambda_i-\rho),\ldots,\prod_{\substack{1\le i\le r\\ i\ne r}}(\lambda_i-\rho)}_{\mu^r\text{elements}})V^T x,$$

which is a polynomial (thus smooth) function of (ρ, x). $\square$

Note that if all the eigenvalues of A are simple, then (9) yields $z' = \mathrm{adj}(A - \rho I)z$, where adj denotes the adjugate (transposed matrix of cofactors); see [Hüp03, 4.1] for details.

In order to fix ideas, we illustrate on a very simple example how the computation of Z can be carried out. Let $A = \left[\begin{smallmatrix}3&1\\1&3\end{smallmatrix}\right]$ with eigenvalues 2 and 4. Let $\rho = 2 + \epsilon$ with $0 \le \epsilon << 1$ and let $x = \left[\begin{smallmatrix}a\\b\end{smallmatrix}\right]$. We have to solve a system in the form

$$(A-\rho I)\begin{bmatrix}z_1\\z_2\end{bmatrix} = \begin{bmatrix}a\\b\end{bmatrix}. \tag{11}$$

Gauss elimination with row pivoting yields

$$\begin{bmatrix}1 & 1-\epsilon\\ 0 & 2\epsilon+\epsilon^2\end{bmatrix}\begin{bmatrix}z_1\\z_2\end{bmatrix} = \begin{bmatrix}a\\ b-a(1-\epsilon)/(1+\epsilon)\end{bmatrix}. \tag{12}$$

If $\epsilon \ne 0$, the solution is

$$z_\epsilon = \frac{1}{\epsilon(2+\epsilon)}\begin{bmatrix}-(a-b)+\epsilon(a+2\epsilon b)\\ a-b+\epsilon b\end{bmatrix} \tag{13}$$

Now assume $\epsilon = 0$, i.e. ρ_i is an eigenvalue of A. Then (12) becomes

$$\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}z_1\\z_2\end{bmatrix} = \begin{bmatrix}a\\ b-a\end{bmatrix}.$$

If $a \ne b$, then there is no solution. If $a = b$, then there is an infinite number of solutions. Following Algorithm 3.2, we compute $\mathcal{N}(A - \rho I) = (1, -1)^T\mathbb{R}$ and $z = P_{\mathcal{N}(A-\rho I)}(a, b)^T = \frac{a-b}{2}(1, -1)^T := z_0$. Notice that if $x = (a, b)^T$ is close to the eigenvector $(1, -1)^T$, then $a-b \ne 0$ and the span of z_ϵ is close to the span of z_0 when ϵ is small. This illustrates the fact, discussed e.g. in [PW79], that although z_ϵ blows up in a neighbourhood of $\epsilon = 0$, its direction $\mathrm{span}(z_\epsilon)$ is well conditioned. This also shows that $\mathrm{span}(z_0)$ is the continuous extension of $\mathrm{span}(z_\epsilon)$ for $\epsilon \to 0$. The continuity of F_{GRQI} will be discussed in Section 5.

4 Rank deficiency

Let ψ be as in (3) and $\mathcal{RD}$ as in (4). With the notation of Definition 3.1 and Algorithm 3.2, let

$$\psi^0 := \{\mathrm{span}(\tilde{X}) : \exists \rho_i = \lambda_j \text{ and } \dim(P_{\mathcal{E}_j}\mathcal{S}_i) < \dim(\mathcal{S}_i)\} \tag{14}$$

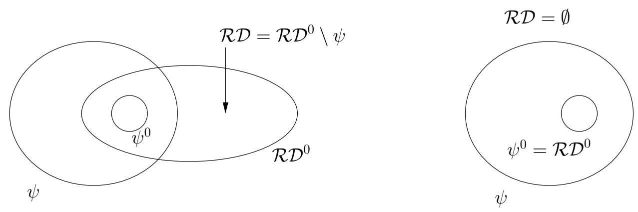

Figure 1: Left: Relations between the sets ψ, ψ^0, $\mathcal{RD}$ and $\mathcal{RD}^0$ defined as in (3), (14), (4) and (15), respectively. Right: Same, assuming that $\mathcal{RD}^0 = \psi^0$.

be the set of all $\mathcal{S}$'s for which the second line of (7) is invoked, and let

$$\mathcal{RD}^0 := \{\mathcal{S} : \dim(F_{\mathrm{GRQI}}(\mathcal{S})) < p\} \tag{15}$$

be the set of all $\mathcal{S}$'s for which the extended GRQI mapping produces a rank reduction. It directly comes that $\psi^0 \subseteq \psi$, $\mathcal{RD} = \mathcal{RD}^0 \setminus \psi$ and $\psi^0 \subseteq \mathcal{RD}^0$. The situation is illustrated on the left-hand side of Figure 1. In the present section we give numerical evidence that $\mathcal{RD}^0 = \psi^0$ (Conjecture 4.2) and we prove that it holds true for $p = 1$ and $p = 2$ (Proposition 4.1). The situation when $\mathcal{RD}^0 = \psi^0$ is illustrated on the right-hand side of Figure 1. Obviously from Figure 1, Conjecture 4.2 ($\mathcal{RD}^0 = \psi^0$) is stronger than Conjecture 2.1 ($\mathcal{RD} = \emptyset$).

Note that $\dim(P_{\mathcal{E}_j}\mathcal{S}_i) < \dim(\mathcal{S}_i)$ if and only if $\mathcal{S}_i$ and $\mathcal{E}_j$ contain orthogonal vectors, or equivalently have a principal angle of $\frac{\pi}{2}$, which always happens if $\dim(\mathcal{E}_j) < \dim(\mathcal{S}_i)$. It is also worthwhile to note that because they can be written as root sets of polynomials, ψ^0 and $\mathcal{RD}$ are thin[3] zero-measure subsets of $\mathrm{Grass}(p, n)$.

Proposition 4.1 *Let ψ^0 and $\mathcal{RD}^0$ be as in* (14) *and* (15)*, respectively. If $p \in \{1, 2\}$ then $\mathcal{RD}^0 = \psi^0$.*

Proof. Since $\psi^0 \subseteq \mathcal{RD}^0$, it is sufficient to show that $\mathcal{RD}^0 \subseteq \psi^0$. This is straightforward to check that this holds when $p = 1$: The matrix Z computed by Algorithm 3.2 has only one column, and if it is zero (i.e. rank-deficient) then necessarily $\rho(X)$ is an eigenvalue λ_i of A and X is orthogonal to the corresponding invariant subspace $\mathcal{E}_i$.

The property $\mathcal{RD} \subseteq \psi^0$ also holds when $p = 2$, as we now show. Let $p = 2$ and let X, ρ_1, ρ_2 and Z be as in Algorithm 3.2. Three cases have to be considered.
Case I: $\rho_1 \notin \mathrm{spec}(A)$ and $\rho_2 \notin \mathrm{spec}(A)$. Then $Z = [Z_1|Z_2]$ verifies

$$(A - \rho_1 I)Z_1 = X_1 \tag{16a}$$

$$(A - \rho_2 I)Z_2 = X_2. \tag{16b}$$

Suppose that $X \in \mathcal{RD}^0$, i.e. Z is rank deficient. Then, since neither Z_1 nor Z_2 vanishes (otherwise X_1 or X_2 would vanish), one has $Z_2 = \mu Z_1$ with $\mu \neq 0$. Left-multiply (16a)

[3] We say that a subset G of a topological set S is a "generic subset" of S if the interior of G is dense in S. The complement of a generic subset is called a "thin subset": its closure has no interior point. Intuitively, a generic subset G is "robust": If x does not belong to G, then there are points arbitrarily close to x that belong to the interior of G. For example, the set of irrational numbers in $\mathbb{R}$ is not a generic subset, although it has full measure.

and (16b) by $(A - \rho_2 I)$ and $(A - \rho_1 I)$, respectively, and use the fact that $(A - \rho_2 I)$ and $(A - \rho_1 I)$ commute to obtain

$$\mu(A - \rho_2 I)X_1 = (A - \rho_1 I)X_2.$$

Left-multiplying this equation by X_1^T yields $\rho_1 - \rho_2 = 0$. Then from (16) one obtains $X_2 = \mu X_1$, a contradiction because X_1 and X_2 are linearly independent. Hence, in this Case I, $X \notin \mathcal{RD}^0$ and the proposition holds.
Case II: One and only one of the ρ's is an eigenvalue of A. Without loss of generality, assume $\rho_1 \in \operatorname{spec}(A)$ and $\rho_2 \notin \operatorname{spec}(A)$. Two sub-cases have to be considered: $Z_1 = 0$ and $Z_1 \neq 0$. In the first case, $X \in \mathcal{RD}^0$ and $X \in \psi^0$, so the proposition holds. In the second case we show that $X \notin \mathcal{RD}^0$, so the proposition also holds. Algorithm 3.2 implies that Z_1 is an eigenvector of A with eigenvalue ρ_1, i.e. $AZ_1 = \rho_1 Z_1$. Suppose that $X \in \mathcal{RD}^0$, i.e. Z is rank deficient, i.e. $Z_2 = \mu Z_1$. Then one has $(A - \rho_2 I)\mu Z_1 = X_2$, and thus $(\rho_1 - \rho_2)\mu Z_1 = X_2$. Hence X_2 is, like Z_1, an eigenvector of A, and thus $\rho_2 \in \operatorname{spec}(A)$. This is a contradiction.
Case III: Both ρ_1 and ρ_2 are eigenvalues of A. Then it is direct from the definitions of Algorithm 3.2, ψ^0 and $\mathcal{RD}^0$ that if $X \in \mathcal{RD}^0$ then $X \in \psi^0$. □

For the case $p \geq 3$, the same result can only be conjectured.

Conjecture 4.2 *Let $\mathcal{RD}^0$ and ψ^0 be defined as in* (15) *and* (14) *with notations introduced in Section 3. Then $\mathcal{RD}^0 = \psi^0$ holds true for all p.*

This conjecture says that Z in Algorithm 3.2 is rank-deficient if and only if there is a loss of rank in at least one of the operations $Z_i = P_{\mathcal{E}_j} X_i$.

We now present a numerical experiment supporting Conjecture 4.2. In order to assess the "distance to rank-deficiency" of Z, we define a smooth real-valued function $\operatorname{aper}(Z)$ measuring the "aperture" of the columns of Z:

$$\operatorname{aper}(Z) = \operatorname{vol}(\hat{Z}) = \sqrt{\det(\hat{Z}^T \hat{Z})}$$

where $\hat{Z}$ is Z with normalized columns and $\operatorname{vol}(\hat{Z})$ denotes the volume spanned by the columns of $\hat{Z}$ [MBI92]. See [MBI92] for an interpretation of the volume of a matrix in terms of principal angles. The justification for normalizing the columns of Z comes from the fact that in the classical RQI the norm of z is ill-conditioned in the neighbourhood of the fixed points while its direction is well-conditioned [PW79]. Note that Z is rank deficient if and only if $\operatorname{aper}(Z) = 0$ or a column of Z is zero, and $\operatorname{aper}(Z) = 1$ when the columns of Z are orthogonal to each other. We chose a diagonal matrix A without repeated elements and we computed the value of $\operatorname{aper}(Z)$ for randomly chosen $\mathcal{S}$'s. The results, displayed on Figure 2, show that $\operatorname{aper}(Z)$ is close to zero only when two Ritz values are close to one same eigenvalue of A. In the limit case where two Ritz values are equal to one same eigenvalue of A, then $\mathcal{S}$ belongs to ψ^0. This suggests that if Z computed by Algorithm 3.2 is rank-deficient then $\mathcal{S}$ belongs to ψ^0.

In conclusion, the result $\mathcal{RD}^0 = \psi^0$ is proven for $p \in \{1, 2\}$ and suggested by our numerical experiments for $p \geq 3$.

Note that in some low dimensional cases, the set ψ^0 reduces to a single point. When $p = 1, n = 3$, ψ^0 reduces to the one-dimensional subspace spanned by

$$\begin{bmatrix} \sqrt{\lambda_3 - \lambda_2} \\ 0 \\ \sqrt{\lambda_2 - \lambda_1} \end{bmatrix}.$$

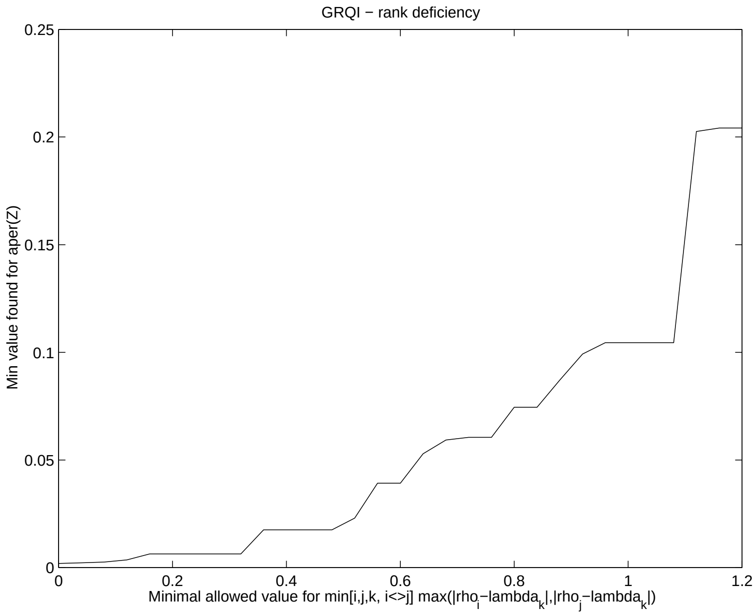

Figure 2: $\text{aper}(Z)$, i.e. the "distance" of Z to rank-deficiency, versus $\min_{i,j,k,i\neq j} \max(|\rho_i - \lambda_k|, |\rho_j - \lambda_k|)$, i.e. the shortest distance between any eigenvalue of A and any cluster of two Ritz values. All data points appear above the displayed curve.

In the case $p = 2, n = 3$, ψ^0 reduces to the two-dimensional subspace spanned by

$$\begin{bmatrix} 0 & \sqrt{\lambda_3 - \lambda_2} \\ 1 & 0 \\ 0 & \sqrt{\lambda_2 - \lambda_1} \end{bmatrix}.$$

This concludes the study of the singularities of the Grassmann-RQI.

5 Smoothness

We now study the smoothness of the mapping $F_{\text{GRQI}} : \mathcal{S} \mapsto \mathcal{S}^+$ given in Definition 3.1. If the mapping is smooth at a point $\mathcal{S}$, then small perturbations on $\mathcal{S}$ will produce small perturbations on $F_{\text{GRQI}}(\mathcal{S})$; more precisely, there exist $c > 0$ and $\epsilon > 0$ such that $\text{dist}(F_{\text{GRQI}}(\mathcal{S}), F_{\text{GRQI}}(\tilde{\mathcal{S}})) \leq c\, \text{dist}(\mathcal{S}, \tilde{\mathcal{S}})$ for all $\tilde{\mathcal{S}}$ such that $\text{dist}(\mathcal{S}, \tilde{\mathcal{S}}) \leq \epsilon$. In the $p = 1$ case (RQI), the mapping F_{GRQI} is smooth on $\text{Grass}(p, n) \setminus \psi^0$, and it cannot be continuously extended on ψ^0 [BS89]. Moreover, as shown in the previous section, $\psi^0 = \mathcal{RD}^0$ when $p \in \{1, 2\}$; see also Conjecture 4.2. For general p, we have the following result stating that F_{GRQI} is smooth everywhere except at some specific points.

Theorem 5.1 *Let $\hat{\mathcal{S}} \in \text{Grass}(p, n)$ be such that $\hat{\mathcal{S}}^+ := F_{\text{GRQI}}(\hat{\mathcal{S}})$ is p-dimensional, where F_{GRQI} is the extended Grassmann-RQI mapping defined in Definition 3.1 (in other words, $\hat{\mathcal{S}}$ does not belong to $\mathcal{RD}^0$). Then F_{GRQI} is continuous in a neighbourhood of $\hat{\mathcal{S}}$.*
Moreover, let $\hat{X}$ be orthonormal with $\text{span}(\hat{X}) = \hat{\mathcal{S}}$ and assume that

(A1) there is no multiple eigenvalue of $\hat{X}^T A\hat{X}$ that is also an eigenvalue of A.
Then F_{GRQI} is smooth (C^∞) in a neighbourhood of $\hat{\mathcal{S}}$.

Proof. The principle of the proof is to decompose F_{GRQI} into a succession of continuous, resp. smooth, operations. A difficulty arises because an extraction of Ritz values is involved in the way F_{GRQI} removes singularities of the first kind (see Section 2 for singularities). The extraction of Ritz values cannot in general be defined smoothly unless they are simple, and this is why we require assumption (A1) to prove smoothness of F_{GRQI} around $\hat{\mathcal{S}}$.

In a neighbourhood of $\hat{\mathcal{S}}$, the mapping F_{GRQI} can be decomposed into the following sequence of operations

$$\begin{aligned}&\mathcal{S} \mapsto Y := \sigma_{\hat{X}}\mathcal{S} \mapsto X := Y(I_p + Y^TY)^{-1/2} \mapsto (x_1,\dots,x_p,\rho_1,\dots,\rho_q)\\ &\mapsto (\alpha_{\rho_1} z_{\rho_1,x_1}|\dots|\alpha_{\rho_q} z_{\rho_q,x_q}|Z_{(x_{q+1},\dots,x_p)}) \mapsto \mathcal{S}^+ := \mathrm{span}(\alpha_{\rho_1} z_{\rho_1,x_1}|\dots|\alpha_{\rho_q} z_{\rho_q,x_q}|Z_{(x_{q+1},\dots,x_p)}).\end{aligned} \tag{17}$$

In the rest of the proof we define these operations and study their smoothness properties.

In the first operation $\mathcal{S} \mapsto Y$, σ denotes the cross section mapping of $\mathrm{Grass}(p,n) = \mathrm{ST}(p,n)/\mathrm{GL}_p$ defined by $\sigma_W \mathrm{span}(Y) := Y(W^TY)^{-1}W^TW$, see [AMS04]. The operation $\mathcal{S} \mapsto Y$ is smooth around $\hat{\mathcal{S}}$ in view of the definition of the differentiable structure of the Grassmann manifold; see e.g. [HM94, Section C.4] or [AMS04].

The second operation $Y \mapsto X$ is smooth. Observe that the composed operation $\mathcal{S} \mapsto X$ smoothly assigns an orthonormal basis X to each subspace $\mathcal{S}$ in a neighbourhood of $\hat{\mathcal{S}}$.

The third operation is required to satisfy the following conditions: $(x_1,\dots,x_p)(X)$ is an orthonormal basis of the span of X and

$$[(x_1,\dots,x_p)(X)]^T\ A\ (x_1,\dots,x_p)(X) = \mathrm{diag}(\rho_1(X),\dots,\rho_q(X),B(X))$$

where $\rho_1(\hat{X}),\dots,\rho_q(\hat{X}) \in \mathrm{spec}(A)$ and $\mathrm{spec}(B(\hat{X})) \cap \mathrm{spec}(A) = \emptyset$. It is always possible to make this operation continuous; see, e.g., [Sun90, Section 3.1]. Under assumption (A1), which guarantees that $\rho_1(\hat{X}),\dots,\rho_q(\hat{X})$ are all simple, it is even possible to make this operation smooth around $\hat{X}$; see, e.g., [Sun90, Theorem 2.1]. The reason for imposing (A1) is that the presence of multiple values may make it impossible to define the operation smoothly; see the counter-example in [ACL93, p. 906].

In the fourth operation, α_{ρ_i} is defined as in (10), z_{ρ_i,x_i} is defined as in (8), and Z_Y denotes the solution of $AZ - ZY^TAY = Y$. In view of Lemma 3.3, the fourth operation is smooth in a neighbourhood of the values corresponding to $\hat{\mathcal{S}}$.

Finally the fifth operation is smooth around the values corresponding to $\hat{\mathcal{S}}$ since $\hat{\mathcal{S}}^+$ is assumed to be p-dimensional.

The conclusion comes from the fact that the composition of continuous, resp. smooth, operations is itself continuous, resp. smooth. $\square$

6 Global convergence

The first step in studying the global convergence of the extended GRQI is to characterize the fixed points of F_{GRQI}. It is clear from Definition 3.1 that the invariant subspaces of A are fixed points of F_{GRQI}. We now show that the converse also holds.

Theorem 6.1 (fixed points) *The fixed points of F_{GRQI} (Definition 3.1) are the p-dimensional invariant subspaces of A.*

Proof. It is directly verified from Definition 3.1 that the p-dimensional invariant subspaces of A are fixed points of F_{GRQI}. It remains to prove that the fixed points of F_{GRQI} are invariant subspaces of A. To this end, let $\mathcal{S}$ be a fixed point under Algorithm 3.2, i.e. $\mathcal{S}^+ = \mathcal{S}$. Let $X = (x_1|\ldots|x_p)$, $\rho_1, \ldots, \rho_s$ and $\mathcal{S}_1, \ldots, \mathcal{S}_s$ be defined as in Algorithm 3.2. Suppose that one of the ρ's, say ρ_i, is not an eigenvalue of A. Then $(A - \rho_i I)$ is invertible. Pick j such that x_j belongs $\mathcal{S}_i$. Then $z = (A - \rho_i I)^{-1} x_j$ belongs to $\mathcal{S}_j^+$, i.e. to $\mathcal{S}$. Therefore

$$z = P_{\mathcal{S}} z = x_1 x_1^T z + \ldots + x_p x_p^T z.$$

Now, pre-multiplying the above equation by $x_j^T(A - \rho_i I)$ yields the contradiction $1 = 0$. So, the ρ's are eigenvalues of A. This means that $\mathcal{S}^+$ (equal to $\mathcal{S}$) is spanned by eigenvectors of A by construction, which means that $\mathcal{S} = \mathcal{S}^+$ is an invariant subspace of A. $\square$

Corollary 6.2 *Assuming that the eigenvalues of A are distinct, the extended GRQI mapping (Definition 3.1) is smooth at its fixed points.*

Proof. This is a consequence of Theorem 5.1 and Theorem 6.1. $\square$

We now consider the periodic orbits of the iteration. In the classical $p = 1$ case (RQI), the periodic orbits have been completely characterized. Let e_i denote the ith coordinate vector and let $f_{ij} = e_i + e_j$. Consider the extended RQI mapping F_{RQI}, assuming as above —without loss of generality— that A is diagonal. Assume moreover that the eigenvalues of A are distinct. Then each $\text{span}(f_{ij})$ is a point of period 2. Moreover, beside the fixed points $\text{span}(e_i)$, $i = 1, \ldots, n$, there are no other periodic points. More details can be found in [Par74, Par80, BS89]. It follows that for the extended GRQI mapping F_{GRQI}, the p-dimensional subspaces of the form $\text{span}(e_{k_1}, \ldots, e_{k_{p-1}}, f_{ij})$ are periodic of period 2. Whether these are the only possible periodic orbits is not known. The characterization of the periodic orbits for the $p = 1$ case relies on the "monotonic residuals" property [Par74, Par80]. As we will show below, this property does not extend to $p > 1$.

The global convergence of the iteration is well understood in the case $p = 1$. It is shown in [Par74, Par80] that RQI converges either to an eigenvector or to a periodic orbit (of period 2). The proof is based on the following fact. Let $\{x_k\}$ be a sequence of iterates of RQI normalized such that. Define the residual

$$r(x) := \|\Pi A x\| = \|Ax - \rho(x)x\| \tag{18}$$

where $\|\cdot\|$ denotes the Euclidean norm. Then $r(x_{k+1}) \leq r(x_k)$ for all k.

The two classical generalizations of (18) for X orthonormal $n \times p$ are

$$r_2(X) := \|\Pi A X\|_2 = \|AX - XX^T AX\|_2 = \sigma_{max}(\Pi AX) \tag{19}$$

and

$$r_F(X) := \|\Pi AX\|_F = \sqrt{\text{trace}(XA\Pi AX)}, \tag{20}$$

where σ_{max} denotes the maximal singular value and $\Pi := (I - XX^T)$ is the projector into the orthogonal complement of the span of X. It is easily checked that both definitions only depend on the span of X orthonormal. It turns out that neither r_2 nor r_F systematically decreases under GRQI, as the following counter-example shows.[4] Take $A = \text{diag}(1, 2, 3, 4)$

[4]Interestingly, in the particular case $n = 3$, $p = 2$, no increase in r_2 and r_F has been observed in experiments.

and

$$\mathcal{S} = \text{span}\begin{bmatrix} -0.2335 & 0.3921 \\ -0.6426 & -0.2729 \\ -0.4499 & 0.7565 \\ -0.5745 & -0.4466 \end{bmatrix},$$

then $r_2(\mathcal{S}^+) - r_2(\mathcal{S}) = 0.4053$ and $r_F(\mathcal{S}^+) - r_F(\mathcal{S}) = 0.1299$. This compromizes the possibility of pursuing a global convergence analysis of GRQI along the lines of Parlett and Kahan's proof [Par80].

7 Conclusion and open questions

We have presented results pertaining to the global behaviour of the GRQI method. An extended GRQI mapping has been proposed. It is defined everywhere on the Grassmann manifold and is identical to the original GRQI mapping wherever the latter is defined. The invariant subspaces of A coincide with the fixed points of the extended mapping. Moreover, under generic conditions on A, the extended mapping is smooth on a generic full-measure subset of the Grassmann manifold.

Several questions remain open. In the case $p = 1$ and $p = 2$, we have characterized the subspaces $\mathcal{Y}$ for which $F_{\text{GRQI}}(\mathcal{Y})$ is rank-deficient. For $p \geq 3$ the result is conjectured and supported by a dedicated numerical experiment; see also the weaker (but easier to formulate) Conjecture 2.1 about rank-deficiency of Z in the equation $AZ - ZX^TAX = X$. In the case $p = 1$, Batterson and Smillie [BS89] show that there is no point in ψ^0 at which the RQI mapping admits a continuous extension; whether a similar result holds true for general p is an open question. In the case $p = 1$, Parlett and Kahan have proven that the RQI method converges either to an eigenvector of A or to a periodic orbit of period 2 (see [Par74, Par80]). The technique of proof does not straightforwardly extend to $p > 1$: it is based on a result (monotonic decrease of the residual) that does not hold in general for $p > 1$. Nevertheless, numerical experiments reported in [ASVM04] show systematic convergence to the invariant subspaces of A.

Acknowledgements

The authors would like to thank R. Sepulchre and P. Van Dooren for helpful comments.

References

[Abs03] P.-A. Absil, *Invariant subspace computation: A geometric approach*, Ph.D. thesis, Faculté des Sciences Appliquées, Université de Liège, Secrétariat de la FSA, Chemin des Chevreuils 1 (Bât. B52), 4000 Liège, Belgium, 2003.

[ACL93] A. L. Andrew, K.-w. E. Chu, and P. Lancaster, *Derivatives of eigenvalues and eigenvectors of matrix functions*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 4, 903–926.

[AMS04] P.-A. Absil, R. Mahony, and R. Sepulchre, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, Acta Appl. Math. **80** (2004), no. 2, 199–220.

[AMSV02] P.-A. Absil, R. Mahony, R. Sepulchre, and P. Van Dooren, *A Grassmann-Rayleigh quotient iteration for computing invariant subspaces*, SIAM Review **44** (2002), no. 1, 57–73.

[ASVM04] P.-A. Absil, R. Sepulchre, P. Van Dooren, and R. Mahony, *Cubically convergent iterations for invariant subspace computation*, to appear in SIAM J. Matrix Anal. Appl., 2004.

[BS89] S. Batterson and J. Smillie, *The dynamics of Rayleigh quotient iteration*, SIAM J. Numer. Anal. **26** (1989), no. 3, 624–636.

[Cra51] S. H. Crandall, *Iterative procedures related to relaxation methods for eigenvalue problems*, Proc. Roy. Soc. London **207** (1951), no. 1090, 416–423.

[dSB81] Eurice de Souza and S. P. Bhattacharyya, *Controllability, observability and the solution of $AX - XB = C$*, Linear Algebra Appl. **39** (1981), 167–188. MR 82h:15021

[ES99] L. Eldén and V. Simoncini, *Inexact Rayleigh quotient-type methods for subspace tracking*, Tech. Report 1172, Istituto di Analisi Numerica del CNR, December 1999.

[Hea77] John Z. Hearon, *Nonsingular solutions of $TA - BT = C$*, Linear Algebra and Appl. **16** (1977), no. 1, 57–63. MR 56 #15676

[HM94] U. Helmke and J. B. Moore, *Optimization and dynamical systems*, Springer, 1994.

[Hüp02] K. Hüper, *A calculus approach to matrix eigenvalue algorithms*, Habilitation Dissertation, July 2002, Mathematisches Institut, Universität Würzburg, Germany.

[Hüp03] ———, *A dynamical system approach to matrix eigenvalue algorithms*, Mathematical Systems Theory in Biology, Communications, Computation (J. Rosenthal and D. S. Gilliam, eds.), The IMA Volumes in Mathematics and its Applications, vol. 134, Springer Verlag, New York, 2003, pp. 257–274.

[MBI92] J. Miao and A. Ben-Israel, *On principal angles between subspaces in $\mathbb{R}^n$*, Linear Algebra Appl. **171** (1992), 81–98.

[Ost59] A. M. Ostrowski, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. I-VI*, Arch. Rational Mech. Anal. (1959), 1, 233–241, 2, 423–428, 3, 325–340, 3, 341–347, 3, 472–481, and 4, 153–165.

[Par74] B. N. Parlett, *The Rayleigh quotient iteration and some generalizations for nonnormal matrices*, Mathematics of Computation **28** (1974), no. 127, 679–693.

[Par80] ———, *The symmetric eigenvalue problem*, Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632, 1980, republished by SIAM, Philadelphia, 1998.

[PS95] R. D. Pantazis and D. B. Szyld, *Regions of convergence of the Rayleigh quotient iteration method*, Numer. Linear Algebra Appl. **2** (1995), no. 3, 251–269.

[PW79] G. Peters and J. H. Wilkinson, *Inverse iteration, ill-conditioned equations and Newton's method*, SIAM Review **21** (1979), no. 3, 339–360.

[RR02] A. C. M. Ran and L. Rodman, *A class of robustness problems in matrix analysis*, Interpolation Theory, Systems Theory and Related Topics, The Harry Dym Anniversary Volume (D. Alpay, I. Gohberg, and V. Vinnikov, eds.), Operator Theory: Advances and Applications, vol. 134, Birkhäuser, 2002, pp. 337–383.

[Smi97] P. Smit, *Numerical analysis of eigenvalue algorithms based on subspace iterations*, Ph.D. thesis, CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, 1997.

[SS90] G. W. Stewart and J. G. Sun, *Matrix perturbation theory*, Academic Press, 1990.

[Sun90] J.-g. Sun, *Multiple eigenvalue sensitivity analysis*, Linear Algebra Appl. **137/138** (1990), 183–211.

[Wim88] Harald K. Wimmer, *Linear matrix equations, controllability and observability, and the rank of solutions*, SIAM J. Matrix Anal. Appl. **9** (1988), no. 4, 570–578. MR 90b:15012