A DIFFERENTIAL-GEOMETRIC LOOK AT THE JACOBI–DAVIDSON FRAMEWORK

PIERRE-ANTOINE ABSIL* AND MICHIEL E. HOCHSTENBACH†

Abstract. The problem of computing a p-dimensional invariant subspace of a symmetric positive-definite matrix pencil of dimension n is interpreted as computing a zero of a tangent vector field on the Grassmann manifold of p-planes in $\mathbb{R}^n$. The theory of Newton's method on manifolds is applied to this problem, and the resulting Newton equations are interpreted as block versions of the Jacobi–Davidson correction equation for the generalized eigenvalue problem.

1. Introduction. The Jacobi–Davidson method (JD) [32] is a method to compute certain eigenpairs of standard or generalized eigenvalue problems. JD has been particularly successful for standard eigenproblems where interior eigenvalues are required, and for generalized types of eigenproblems. JD belongs to the the class of subspace methods, where low-dimensional subspaces are exploited to find approximations to sought eigenvectors. In line with many other subspace methods, JD has two main stages:

(i) The subspace extraction, where approximate eigenpairs are determined from a given search space. This is often done by the Rayleigh–Ritz method, or variants such as the harmonic Rayleigh–Ritz approach (see, e.g., [36]).

(ii) The subspace expansion, where the search space is expanded with an (inexact) solution to the so-called correction equation.

We refer to [20] for a recent overview of several aspects of the JD method; see also [22].

In [33, §6], JD for the standard eigenvalue problem is interpreted as a Newton method. The interpretation is readily extended to the generalized eigenvalue problem as follows. Let (A, B) be a symmetric positive-definite matrix pencil; we refer to Section 3 for the necessary background on the generalized eigenvalue problem. We are interested in an eigenvector y of (A, B). Let $\widetilde{u}, w \in \mathbb{R}^n$ be fixed for the time being. In order to remove the scale indeterminacy of eigenvectors, we impose the normalization $\widetilde{u}^T y = 1$. Consider the function defined for all $u \in \{u : \widetilde{u}^T u = 1\}$ by

$$F(u) = Au - \theta Bu \quad \text{with} \quad \theta = \theta(u) = \frac{w^T Au}{w^T Bu},$$

where we assume that $w^T Bu \neq 0$. Function F maps the hyperplane $\{u : \widetilde{u}^T u = 1\}$ to the hyperplane $w^\perp$. Observe that u with $\widetilde{u}^T u = 1$ is an eigenvector of (A, B) if and only if $F(u) = 0$. If u is the current approximation, then the next Newton iterate for F is $u + s$, where $s \perp \widetilde{u}$ satisfies

$$(\mathrm{D}F(u))s = -F(u). \tag{1.1}$$

It may be checked that the Jacobian of F acting on $\widetilde{u}^\perp$ is given by

$$(\mathrm{D}F(u))s = \left(I - \frac{Buw^T}{w^T Bu}\right)(A - \theta(u)B)s \quad \text{for} \quad s \perp \widetilde{u},$$

*Université catholique de Louvain, Belgium, sites.uclouvain.be/absil/

†Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB, The Netherlands, www.win.tue.nl/∼hochsten.

hence the Newton equation (1.1) reads

$$\left(I - \frac{Buw^T}{w^TBu}\right)(A - \theta(u)B)s = -(A - \theta(u)B)u \quad \text{for} \quad s \perp \widetilde{u}. \tag{1.2}$$

This Newton process converges locally quadratically to an eigenvector y with $\widetilde{u}^T y = 1$, for fixed $\widetilde{u}$ and w, provided that $w^TBy \neq 0$. However, the adaptive choice $\widetilde{u} = u$ and $w = u$ also leads to locally quadratic convergence. For this choice, the Newton correction equation (1.2) is precisely the correction equation that appears in JD.

In [26], a block Newton method was given for the standard eigenvalue problem, and a connection was made with JD in the appendix. Expected advantages of a block method are better robustness or efficiency in the presence of clustered eigenvalues, as well as exploiting higher-level BLAS; see, e.g., the discussion in [12, §5.1.4], where a block JD for the generalized eigenvalue problem is outlined. In particular, if the desired eigenvalues are multiple or clustered, then difficulties can arise in the (non-block) Jacobi–Davidson method, because the Jacobi correction equation (1.1) becomes ill conditioned. Resorting to a block version allows one to "split" the spectrum at a wider eigenvalue gap.

In this paper, we obtain a class of block Jacobi correction equations for the generalized eigenvalue problem. Our approach consists in characterizing the p-dimensional invariant subspaces of (A, B) as the zeros of a tangent vector field on the Grassmann manifold of p-planes in $\mathbb{R}^n$. The Grassmann manifold is described as a quotient of the set $\mathbb{R}_*^{n\times p}$ of all $n \times p$ matrices of full (column) rank, where the equivalences classes gather all matrices that have the same column space. By applying Shub's manifold-based Newton method [30, §3] to the tangent vector field, and by exploiting the leeway offered by the quotient geometry framework, we obtain a whole class of block Jacobi correction equations for the generalized eigenvalue problem.

The paper is organized as follows. Section 2 gives a brief overview of algorithms on manifolds in connection with the eigenvalue problem. The generalized eigenvalue problem is described in Section 3. The tangent vector field on the Grassmann manifold is introduced in Section 4. The geometric Newton method for the vector field is worked out in Section 5. Connections with JD are established in Section 6.

The forthcoming developments will make use of differential-geometric objects on the Grassmann manifold viewed as a quotient of $\mathbb{R}_*^{n\times p}$. However, the necessary differential-geometric concepts are quite limited, and this paper is meant to be accessible without any background in differential geometry.

2. Algorithms on manifolds and eigenvalue problems. The best known field of application of differential geometry is probably relativity theory. More surprisingly perhaps, techniques of differential and Riemannian geometry have found applications in several areas of science and engineering (such as crystallography [25], thermodynamics [10], and information theory [6, 35]), and in particular, they have been utilized to design and analyze eigenvalue algorithms. Numerous papers belong to this line of research, including [7, 30, 17, 18, 34, 21, 27, 11, 15, 19, 23, 26, 3, 24, 29, 16, 8, 1, 9].

That the Jacobi–Davidson approach, including a block version thereof, is closely related to Newton's method on Grassmann manifolds, was pointed out in [26]. Since then, the area of numerical computations on manifolds has made progress in several directions, some of which will be exploited in this work. Whereas the seminal papers [34, 11] were systematically making use of the Riemannian connection and the

Riemannian exponential, more recent papers [28, 5, 1, 4] have relied on the concepts of *retraction* and of *locally smooth family of parameterizations* to relax the Riemannian exponential into a broader class of mappings that offer opportunities to reduce the numerical burden while preserving convergence properties. Likewise, the Newton method on manifolds stated and analyzed in [30, 4] allows for using of any affine connection, instead of restricting to the Riemannian one as in [34, 11]. As a consequence, the Newton method on Riemannian manifolds stated in [34] has turned, without losing its quadratic local convergence, into a class of geometric Newton methods which vary according to the choice of an affine connection and of a retraction.

3. The generalized eigenvalue problem: notation and assumptions. Given two $n \times n$ matrices A and B, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue*, that $u \in \mathbb{R}^n \backslash \{0\}$ is an *eigenvector*, and that (λ, u) is an *eigenpair* of the *pencil* (A, B) if

$$Ax = \lambda Bx.$$

Finding eigenpairs of a matrix pencil is known as the *generalized eigenvalue problem*. From now on, we assume that A is symmetric and B is symmetric positive-definite (i.e., $x^T Bx > 0$ for all $x \neq 0$); the pencil (A, B) and the associated generalized eigenvalue problem are then termed *symmetric positive-definite*, abbreviated *S/PD* [36, §4.1]. It follows that the eigenvalues of the pencil are all real and the eigenvectors can be chosen to form a B-orthonormal basis. A subspace $\mathcal{Y}$ is a *(generalized) invariant subspace* (or a *deflating subspace*) [14, §7.7.8] of the S/PD pencil (A, B) if $B^{-1}Ay \in \mathcal{Y}$ for all $y \in \mathcal{Y}$; this can also be written $B^{-1}A\mathcal{Y} \subseteq \mathcal{Y}$ or $A\mathcal{Y} \subseteq B\mathcal{Y}$. It is readily seen that $\mathcal{Y}$ is a one-dimensional invariant subspace of (A, B) if and only if $\mathcal{Y}$ is spanned by an eigenvector of (A, B). More generally, every invariant subspace of an S/PD pencil is spanned by eigenvectors of (A, B). Clearly, the generalized eigenvalue problem reduces to the standard eigenvalue problem when $B = I$.

Given an integer $1 \leq p \leq n$, we let $\mathbb{R}_*^{n \times p}$ denote the set of all $n \times p$ matrices of full (column) rank, and we let $\mathsf{col}(Y)$, termed the *column space* of Y, denote the p-dimensional subspace of $\mathbb{R}^n$ spanned by the columns of $Y \in \mathbb{R}_*^{n \times p}$, i.e.,

$$\mathsf{col}(Y) = \{Yw : w \in \mathbb{R}^p\}.$$

The set of all matrices $\hat{Y}$ such that $\mathsf{col}(\hat{Y}) = \mathsf{col}(Y)$ is

$$[Y] := Y\,\mathrm{GL}_p := \{YM : M \in \mathrm{GL}_p\}, \tag{3.1}$$

where

$$\mathrm{GL}_p := \{M \in \mathbb{R}^{p \times p}\}$$

denotes the set of all $p \times p$ invertible matrices. Observe that $\mathcal{Y}$ is a p-dimensional invariant subspace of (A, B) if and only if there is $Y \in \mathbb{R}_*^{n \times p}$ with $\mathcal{Y} = \mathsf{col}(Y)$ such that

$$AY = BYM \tag{3.2}$$

for some $p \times p$ matrix M.

The *multiplicity* of an eigenvalue λ of (A, B) is its multiplicity as a root of the polynomial $\det(A - \lambda B)$. An invariant subspace $\mathsf{col}(Y)$ of (A, B) is termed *simple* [36] or

spectral [13] if the multiplicity of the eigenvalues of M is the same as their multiplicity as eigenvalues of (A, B).

We will let $U \mapsto \widetilde{U}_U$ denote any function on $\mathbb{R}_*^{n\times p}$ into $\mathbb{R}_*^{n\times p}$ that satisfies the following two properties. (i) $\mathsf{col}(\widetilde{U}_U)$ only depends on $\mathsf{col}(U)$, i.e., for all $M \in \mathrm{GL}_p$, there is $N \in \mathrm{GL}_p$ such that $\widetilde{U}_{UM} = \widetilde{U}_U N$. For example, the choice $\widetilde{U} = CU$ for a fixed C is adequate. (ii) For all $U \in \mathbb{R}_*^{n\times p}$, $\widetilde{U}_U^T U$ and $\widetilde{U}_U^T BU$ are invertible. The motivation for imposing invertibility of $\widetilde{U}_U^T BU$ will already become clear in Theorem 4.1. The other assumptions will be instrumental in the differential-geometric approach laid out below.

Finally, we will let

$$P_{E,F} := I - E(F^T E)^{-1} F^T \tag{3.3}$$

denote the projector along the column space of E into the orthogonal complement of the column space of F.

4. Invariant subspaces as zeros of a vector field. We are looking for an iteration function

$$g : \mathbb{R}_*^{n\times p} \to \mathbb{R}_*^{n\times p} \tag{4.1}$$

such that the sequences of iterates $\mathsf{col}(U_k)$ generated by $U_{k+1} = g(U_k)$ converge locally quadratically to the p-dimensional spectral invariant subspaces of (A, B). The quadratic convergence requirement leads us naturally to Newton-type methods.

Moreover, since we are interested in the sequence of subspaces $\mathsf{col}(U_k)$ rather than in the sequence of p-frames U_k, it makes sense to require that g and col commute; in other words, there must be a function G such that $G(\mathsf{col}(U)) = \mathsf{col}(g(U))$ for all $U \in \mathbb{R}_*^{n\times p}$. The domain and the codomain of G are the set $\mathrm{Grass}(p, n)$ of all p-dimensional subspaces in $\mathbb{R}^n$. This set admits a natural manifold structure, as explained in [18, §C.4], and endowed with this structure, it is called a *Grassmann manifold*. The sought Newton-type method will thus be an iteration on the Grassmann manifold $\mathrm{Grass}(p, n)$.

To this end, we will pursue the following strategy. In this section, we will express the problem of computing a p-dimensional invariant subspace of (A, B) as finding a zero of a particular vector field on $\mathrm{Grass}(p, n)$. Then, in Section 5, we will work out Shub's Newton method for this vector field.

The characterization of the p-dimensional invariant subspaces of (A, B) as the zeros of a vector field on $\mathrm{Grass}(p, n)$ relies on the following result.

THEOREM 4.1. *Let $U \in \mathbb{R}_*^{n\times p}$. Under the assumptions of Section 3, $\mathsf{col}(U)$ is an invariant subspace of (A, B) if and only if*

$$AU - BU(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U AU = 0. \tag{4.2}$$

Proof. The "if" part is direct in view of (3.2). For the "only if" part, assume that $\mathsf{col}(U)$ is an invariant subspace of (A, B). Then, in view of (3.2), there is M such that $AU = BUM$. Multiplying this equation on the left by $\widetilde{U}_U^T$ yields that $M = (\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U AU$, hence the claim. □

The rest of this subsection is dedicated to showing that the mapping $U \mapsto AU - BU(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U AU$ that appears in (4.2) represents a vector field on the set of all p-dimensional subspaces of $\mathbb{R}^n$.

The Grassmann manifold $\mathrm{Grass}(p,n)$ can be viewed as the manifold of rank-p symmetric projection operators of $\mathbb{R}^n$; see [16] for details in the context of Newton's method. It can also be viewed as a homogeneous space for the orthogonal group $\mathrm{O}(n)$; see [11]. In the context of this paper, since elements of $\mathrm{Grass}(p,n)$ are represented as column spaces of elements of $\mathbb{R}_*^{n\times p}$, we find it more convenient to rely on the identification of $\mathrm{Grass}(p,n)$ with the quotient space

$$\mathbb{R}_*^{n\times p}/\mathrm{GL}_p = \{[Y] : Y \in \mathbb{R}_*^{n\times p}\},$$

where $[\cdot]$ is as defined in (3.1). The one-to-one correspondence between $\mathrm{Grass}(p,n)$ and $\mathbb{R}_*^{n\times p}/\mathrm{GL}_p$ is given by $\mathsf{col}(Y) \leftrightarrow [Y]$. This identification was mentioned in [18, §C.4] and further exploited in [3, 4]. In view of the identification $\mathrm{Grass}(p,n) \simeq \mathbb{R}_*^{n\times p}/\mathrm{GL}_p$, the sought Newton-like iteration G can thus be viewed as an iteration on $\mathbb{R}_*^{n\times p}/\mathrm{GL}_p$.

We now particularize to $\mathbb{R}_*^{n\times p}/\mathrm{GL}_p$ the framework presented, e.g., in [4, §3.5.8] that allows to represent tangent vectors to $\mathbb{R}_*^{n\times p}/\mathrm{GL}_p$ as $n \times p$ matrices by means of so-called horizontal lifts. The difference with the Grassmann-specific developments in [4, Example 3.6.4] is that we depart from the framework of Riemannian submersions, where the horizontal space is constrained to be the orthogonal complement of the vertical space. This additional freedom allows us to obtain a wider class of Newton equations.

For each $U \in \mathbb{R}_*^{n\times p}$, the *vertical space* $\mathcal{V}_U$ is the tangent space to $[U]$ at U. We have

$$\mathcal{V}_U = \{UM : M \in \mathbb{R}^{p\times p}\}.$$

Intuitively, the vertical space $\mathcal{V}_U$ consists of all the elementary variations of U that preserve the column space.

We choose the *horizontal space* $\mathcal{H}_U$ as the set of all $n \times p$ matrices whose columns are orthogonal to the columns of $\widetilde{U}_U$, i.e.

$$\mathcal{H}_U := \{Z \in \mathbb{R}^{n\times p} : \widetilde{U}_U^T Z = 0\}. \tag{4.3}$$

The horizontal space is required to be transverse to the vertical space, i.e., $\mathcal{H}_U \cap \mathcal{V}_U = \{0\}$; this is equivalent to the condition that $U^T\widetilde{U}_U$ be invertible, which is a standing assumption (see Section 3). Moreover, we require the compatibility condition that $\mathcal{H}_{UM} = \mathcal{H}_U$ for all $M \in \mathrm{GL}_p$; this is ensured by the standing assumption that $\mathsf{col}(\widetilde{U}_U)$ only depends on $\mathsf{col}(U)$ (see Section 3).

The purpose of the horizontal space is to provide a unique matrix representation of elementary variations of p-dimensional subspaces of R^n. Given an elementary variation $\xi_{\mathcal{U}}$ of the column space $\mathcal{U}$ of U, there is in $\mathcal{H}_U$ one and only one elementary variation $\overline{\xi}_U$ of U that has the same effect as $\xi_{\mathcal{U}}$, in the sense that, for all real-valued functions f on $\mathrm{Grass}(p,n)$, it holds that $\mathrm{D}(f \circ \mathsf{col})(U)[\overline{\xi}_U] = \mathrm{D}f(\mathsf{col}(U))[\xi_{\mathcal{U}}]$. In the parlance of differential geometry, $\xi_{\mathcal{U}}$ is a *tangent vector* to $\mathrm{Grass}(p,n)$ at $\mathcal{U}$, and $\overline{\xi}_U$ is the *horizontal lift* of $\xi_{\mathcal{U}}$ at U. The set of all tangent vectors to $\mathrm{Grass}(p,n)$ at $\mathcal{U}$ is called the *tangent space* to $\mathrm{Grass}(p,n)$ at $\mathcal{U}$ and denoted by $T_{\mathcal{U}}\mathrm{Grass}(p,n)$. Observe that $\overline{\xi}_U$, in spite of its somewhat unusual notation and its differential geometric origin, is nothing else than an $n \times p$ real matrix. It can be shown (see [4, Proposition 3.6.1]) that the horizontal lifts at different points U and UM of a same equivalence class $[U]$ satisfy the relation

$$\overline{\xi}_{UM} = \overline{\xi}_U M \tag{4.4}$$

for all $M \in \mathrm{GL}_p$. And any vector field $\mathbb{R}_*^{n\times p} \ni U \mapsto \overline{\xi}_U \in \mathbb{R}^{n\times p}$ that satisfies (4.4) is a bona-fide horizontal lift.

Now, returning to the generalized eigenvalue problem for the S/PD pencil (A, B), consider

$$\overline{\xi}_U := P_{BU,\widetilde{U}_U} AU, \tag{4.5}$$

where $P_{BU,\widetilde{U}_U} = I - BU(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U^T$ in keeping with the notation introduced in (3.3). Recall the standing assumption that $\widetilde{U}_U^T BU$ is invertible; hence the right-hand side of (4.5) is well defined. Moreover, it is readily checked that $\overline{\xi}_U$ of (4.5) satisfies (4.4). Thus $\overline{\xi}_U$, being a horizontal lift, defines a vector field ξ on $\mathrm{Grass}(p, n)$. In view of Theorem 4.1, searching for a p-dimensional invariant subspace of $B^{-1}A$ amounts to searching for a zero of the vector field ξ on $\mathrm{Grass}(p, n)$ defined by the horizontal lift (4.5).

5. Geometric Newton for invariant subspace computation. We now work out the geometric Newton equation for the vector field ξ on $\mathrm{Grass}(p, n)$ defined in the previous section.

The geometric Newton method for computing a zero of a vector field ξ on a manifold $\mathcal{M}$ requires an *affine connection* ∇ on $\mathcal{M}$, which can be thought of as a generalization of the directional derivative; for details, see, e.g., [4, §5.2]. In the next paragraph, we proceed to describe a class of affine connections on $\mathrm{Grass}(p, n)$.

Let $\widetilde{W}_U$ be an $n \times p$ matrix that depends on U in such a way that $\mathsf{col}(\widetilde{W}_U)$ is constant on the equivalence classes $[U]$, and such that $\widetilde{U}_U^T \widetilde{W}_U$ is invertible for all U. Define ∇ by

$$\overline{(\nabla_{\eta_{\mathsf{col}(U)}}\xi)}_U = P_{\widetilde{W}_U,\widetilde{U}_U} \mathrm{D}\overline{\xi}(U)[\overline{\eta}_U], \tag{5.1}$$

where ξ is a vector field on $\mathrm{Grass}(p, n)$ and $\eta_{\mathsf{col}(U)}$ is a tangent vector to $\mathrm{Grass}(p, n)$ at $\mathsf{col}(U)$. Observe that $\nabla_\eta \xi$ is a tangent vector to $\mathrm{Grass}(p, n)$ at $\mathsf{col}(U)$ and that $\overline{(\nabla_{\eta_{\mathsf{col}(U)}}\xi)}_U$ denotes the horizontal lift of that tangent vector. It can be checked that the right-hand side of (5.1) satisfies the compatibility condition (4.4) of horizontal lifts, hence (5.1) is a legitimate definition of a tangent vector $\nabla_{\eta_{\mathsf{col}(U)}}\xi$. It can also be checked that the mapping ∇ thus defined has all the properties of an affine connection.

On an abstract manifold $\mathcal{M}$ equipped with an affine connection ∇, the Newton equation at $x \in \mathcal{M}$ for computing a zero of a vector field ξ reads

$$\nabla_{\eta_x}\xi = -\xi_x. \tag{5.2}$$

Now let $\mathcal{M}$ be the Grassmann manifold $\mathrm{Grass}(p, n) = \mathbb{R}_*^{n\times p}/\mathrm{GL}_p$, let x be $\mathsf{col}(U)$, and consider the choice (4.3) for the horizontal space, the choice (5.1) for the affine connection, and the choice (4.5) for the vector field ξ. Then, replacing the symbol $\overline{\eta}_U$ by Z for simplicity of notation, the horizontal lift at U of the left-hand side of the Newton equation (5.2) becomes

$$P_{\widetilde{W}_U,\widetilde{U}_U} \mathrm{D}\overline{\xi}(U)[Z] = P_{\widetilde{W}_U,\widetilde{U}_U} \left(P_{BU,\widetilde{U}_U} AZ - BZ(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U^T AU - BUE[Z] \right), \tag{5.3}$$

where $E[Z] := \mathrm{D}(U \mapsto (\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U^T AU)(U)[Z]$.

We choose

$$\widetilde{W_U} := BU \tag{5.4}$$

to get rid of the $BUE[Z]$ term. The Newton equation (5.2), in its matrix formulation given by the horizontal lift at U, thus becomes

$$P_{BU,\widetilde{U}_U}(AZ - BZ(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U^T AU) = -P_{BU,\widetilde{U}_U} AU, \tag{5.5a}$$

$$\widetilde{U}_U^T Z = 0. \tag{5.5b}$$

(Recall that $P_{BU,\widetilde{U}_U} = I - BU(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U^T$, and that matrix $\widetilde{U}_U$ can be chosen arbitrarily as a function of U under the condition that $\widetilde{U}_U^T U$ and $\widetilde{U}_U^T BU$ be invertible and that the column space of $\widetilde{U}_U$ be constant along the equivalence classes $[U]$.)

With the retraction on $\mathrm{Grass}(p,n)$ chosen as in [4, Example 4.1.5], it follows from the convergence theory of the geometric Newton method [4, Algorithm 4] that the iteration on $\mathrm{Grass}(p,n)$ defined by

$$\mathsf{col}(U) \mapsto \mathsf{col}(U + Z_U), \tag{5.6}$$

where Z_U denotes the solution of (5.5), converges locally, at least quadratically, to the spectral invariant spaces of $B^{-1}A$.

6. Discussion. The Newton map (5.6) is the G function announced in the beginning of Section 4. It is an iteration on the quotient space $\mathbb{R}_*^{n\times p}/\mathrm{GL}_p$, or equivalently on the Grassmann manifold since $\mathbb{R}_*^{n\times p}/\mathrm{GL}_p \simeq \mathrm{Grass}(p,n)$. In practice, the iteration is realized numerically by an iteration function g as in (4.1), such that $\mathsf{col}(U + Z_U) = \mathsf{col}(g(U))$. Any function g such that $g(U) = (U + Z_U)M_U$, where M_U is $p \times p$ and invertible, is suitable. The freedom in M_U can be exploited to keep the iterates (sufficiently close to) orthonormal.

Let W_U be such that $W_U^T BU$ is invertible. Then, without loss of information, we can multiply (5.5a) on the left by P_{BU,W_U} to obtain

$$P_{BU,W_U}(AZ - BZ(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U^T AU) = -P_{BU,W_U} AU. \tag{6.1}$$

On the other hand, a block generalization of the Jacobi correction equation of [31, Algorithm 3.1], with the hypotheses of [31, Theorem 3.2], would rather be

$$P_{BU,W_U}(AZ - BZ(W_U^T BU)^{-1}W_U^T AU) = -P_{BU,W_U} AU. \tag{6.2}$$

As mentioned in [31, §3.3], when these equations are to be solved with unpreconditioned subspace methods, it is desirable to have $W_U = \widetilde{U}_U$; otherwise the domain space $\mathcal{H}_U$ of the linear map $Z \mapsto P_{BU,W_U}(AZ - BZ(\widetilde{U}_U^T BU)^{-1}\widetilde{U}_U^T AU)$ differs from the image space, which implies that powers cannot be formed. In this case where $W_U = \widetilde{U}_U$, (6.1) and (6.2) coincide.

If BU in $P_{BU,\widetilde{U}_U}$ appearing in (5.5) is replaced by U, then some terms that do not go to zero are neglected in the Jacobian, and one can expect that quadratic convergence is lost. This is confirmed in the $p = 1$ case by the experiments reported in [31, §9.1.1].

A full-blown block JD method for the generalized eigenvalue problem would consist in enhancing the Newton equation (5.5) with a Davidson strategy, where U is selected by a Ritz approximation with respect to a subspace spanned by previous corrections

(see [31, Algorithm 4.2]). If the goal is to compute the p-dimensional invariant subspace, $\mathcal{V}$, assumed to be spectral (see Section 3), corresponding to the smallest (resp. largest) eigenvalues of (A, B), and if the p Ritz vectors corresponding to the smallest (resp. largest) Ritz values are used for the next U, then quadratic convergence is preserved. This follows from [2, Proposition 6.1], where the objective function f is the generalized Rayleigh quotient defined by $f(\mathsf{col}(U)) = \mathrm{tr}((U^TBU)^{-1}U^TAU)$ (resp. its opposite). To see this, observe that $\mathcal{V}$ is the global minimizer of f (see, e.g., [4, Proposition 2.1.1]), that it is nondegenerate since it is assumed to be spectral (see the discussion in [4, §6.5.1]), and that the Newton iteration (5.6), since it converges locally quadratically to $\mathcal{V}$, is a descent iteration for f close enough to $\mathcal{V}$.

REFERENCES

[1] P.-A. Absil, C. G. Baker, and K. A. Gallivan, *Trust-region methods on Riemannian manifolds*, Found. Comput. Math., 7 (2007), pp. 303–330.

[2] P.-A. Absil and K. A. Gallivan, *Accelerated line-search and trust-region methods*, SIAM J. Numer. Anal., 47 (2009), pp. 997–1018.

[3] P.-A. Absil, R. Mahony, and R. Sepulchre, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, Acta Appl. Math., 80 (2004), pp. 199–220.

[4] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.

[5] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Anal., 22 (2002), pp. 359–390.

[6] S.-i. Amari and H. Nagaoka, *Methods of information geometry*, vol. 191 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 2000. Translated from the 1993 Japanese original by Daishi Harada.

[7] G. Ammar and C. Martin, *The geometry of matrix eigenvalue methods*, Acta Appl. Math., 5 (1986), pp. 239–278.

[8] M. Baumann and U. Helmke, *Riemannian subspace tracking algorithms on Grassmann manifolds*, in Proceedings of the 46th IEEE Conference on Decision and Control, 2007.

[9] M. Baumann and U. Helmke, *A time-varying Newton algorithm for adaptive subspace tracking*, Mathematics and Computers in Simulation, 79 (2008), pp. 1324–1345.

[10] M. Chen, *On the geometric structure of thermodynamics*, J. Math. Phys., 40 (1999), pp. 830–837.

[11] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 303–353.

[12] R. Geus, *The Jacobi-Davidson algorithm for solving large sparse symmetric eigenvalue problems with application to the design of accelerator cavities*, PhD thesis, Swiss Federal Institute of Technology Zurich, 2002.

[13] I. Gohberg, P. Lancaster, and L. Rodman, *Invariant subspaces of matrices with applications*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons Inc., New York, 1986. Republished 2006 as *Classics in Applied Mathematics 51* by SIAM, Philadelphia.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, third ed., 1996.

[15] U. Helmke and P. Fuhrmann, *Controllability of matrix eigenvalue algorithms: The inverse power method*, Systems and Control Letters, 41 (2000), pp. 57–66.

[16] U. Helmke, K. Hüper, and J. Trumpf, *Newton's method on Grassmann manifolds*, September 2007. arXiv:0709.2205v2.

[17] U. Helmke and J. Moore, *Singular-value decomposition via gradient and self-equivalent flows*, Linear Algebra and Its Applications, 169 (1992), pp. 223–248.

[18] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Communications and Control Engineering Series, Springer-Verlag London Ltd., London, 1994. With a foreword by R. Brockett.

[19] U. Helmke and F. Wirth, *On controllability of the real shifted inverse power iteration*, Systems

and Control Letters, 43 (2001), pp. 9–23.

[20] M. E. Hochstenbach and Y. Notay, *The Jacobi–Davidson method.*, GAMM Mitteilungen, 29 (2006), pp. 368–382.

[21] K. Hueper, U. Helmke, and J. Moore, *Structure and convergence of conventional Jacobi-type methods minimizing the off-norm function*, in Proceedings of the IEEE Conference on Decision and Control, vol. 2, 1996, pp. 2124–2129.

[22] *The Jacobi–Davidson Gateway.* `www.win.tue.nl/casa/research/topics/jd/`.

[23] J. Jordan and U. Helmke, *Controllability of the QR-algorithm on Hessenberg flags*, in Proceeding of the Fifteenth International Symposium on Mathematical Theory of Network and Systems (MTNS 2002), D. S. Gilliam and J. Rosenthal, eds., 2002.

[24] M. Kleinsteuber, U. Helmke, and K. Hüper, *Jacobi's algorithm on compact lie algebras*, SIAM Journal on Matrix Analysis and Applications, 26 (2005), pp. 42–69.

[25] E. Kröner, *The differential geometry of elementary point and line defects in Bravais crystals*, Internat. J. Theoret. Phys., 29 (1990), pp. 1219–1237.

[26] E. Lundström and L. Eldén, *Adaptive eigenvalue computations using Newton's method on the Grassmann manifold*, SIAM J. Matrix Anal. Appl., 23 (2001/02), pp. 819–839.

[27] R. E. Mahony, U. Helmke, and J. B. Moore, *Gradient algorithms for principal component analysis*, J. Austral. Math. Soc. Ser. B, 37 (1996), pp. 430–450.

[28] J. H. Manton, *Optimization algorithms exploiting unitary constraints*, IEEE Trans. Signal Process., 50 (2002), pp. 635–650.

[29] J. H. Manton, U. Helmke, and I. M. Y. Mareels, *A dual purpose principal and minor component flow*, Systems Control Lett., 54 (2005), pp. 759–769.

[30] M. Shub, *Some remarks on dynamical systems and numerical analysis*, in Proc. VII ELAM., L. Lara-Carrero and J. Lewowicz, eds., Equinoccio, U. Simón Bolívar, Caracas, 1986, pp. 69–92.

[31] G. L. G. Sleijpen, A. G. L. Booten, D. R. Fokkema, and H. A. van der Vorst, *Jacobi–Davidson type methods for generalized eigenproblems and polynomial eigenproblems*, BIT, 36 (1996), pp. 595–633.

[32] G. L. G. Sleijpen and H. A. van der Vorst, *A Jacobi–Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425.

[33] ———, *The Jacobi–Davidson method for eigenvalue problems and its relation with accelerated inexact Newton schemes*, in Iterative Methods in Linear Algebra, II., S. D. Margenov and P. S. Vassilevski, eds., vol. 3 of IMACS Series in Computational and Applied Mathematics, New Brunswick, NJ, U.S.A., 1996, IMACS, pp. 377–389.

[34] S. T. Smith, *Optimization techniques on Riemannian manifolds*, in Hamiltonian and gradient flows, algorithms and control, A. Bloch, ed., vol. 3 of Fields Inst. Commun., Amer. Math. Soc., Providence, RI, 1994, pp. 113–136.

[35] S. T. Smith, *Covariance, subspace, and intrinsic Cramér-Rao bounds*, IEEE Trans. Signal Process., 53 (2005), pp. 1610–1630.

[36] G. W. Stewart, *Matrix Algorithms. Vol. II*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Eigensystems.