Ark. Mat., **48** (2010), 57–77
DOI: 10.1007/s11512-009-0098-0

Matrix subspaces and determinantal hypersurfaces

Marko Huhtanen

Abstract. Nonsingular matrix subspaces can be separated into two categories: by being either invertible, or merely possessing invertible elements. The former class was introduced for factoring matrices into the product of two matrices. With the latter, the problem of characterizing the inverses and related nonlinear matrix geometries arises. For the singular elements there is a natural concept of spectrum defined in terms of determinantal hypersurfaces, linking matrix analysis with algebraic geometry. With this, matrix subspaces and the respective Grassmannians are split into equivalence classes. Conditioning of matrix subspaces is addressed.

1. Introduction

The notion of invertible matrix subspace was introduced in connection with factoring a matrix $A \in \mathbb{C}^{n\times n}$ into the product $A = V_1 V_2$ with the factors constrained to belong to prescribed subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$) [21].(1) The linear structure of an invertible matrix subspace is preserved under inversion. Then, if $\mathcal{V}_2$ is invertible with the inverse $\mathcal{W}$, this factoring problem can be converted into considering

$$AW = V_1, \tag{1}$$

with the (nonzero) elements $W \in \mathcal{W}$ and $V_1 \in \mathcal{V}_1$ regarded both as variables. This is an equivalent task in case A is invertible. Certain other bilinear factorization problems can be treated similarly. For instance, also of practical importance is the Kronecker product representation problem $A = V_1 \otimes V_2$.(2) Motivated by such

Supported by the Academy of Finland.

(1) For operator factorization problems, see [24] and [29]. For their perturbation theory, see [5]. On the factorization of a matrix-valued function, see [25].

(2) With $\mathcal{V}_1 = \mathbb{C}^{n_1\times n_1}$ and $\mathcal{V}_2 = \mathbb{C}^{n_2\times n_2}$ the problem is solved with the singular value decomposition [28]. Then we are dealing with the Segre map. This, however, does not apply as soon as $\mathcal{V}_1$ and $\mathcal{V}_2$ are more complicated matrix subspaces.

factorizations, this paper is concerned with matrix subspaces, their spectra and related nonlinear matrix geometries.

Regardless if the matrix subspace $\mathcal{W}$ is invertible or not, consider (1). The validity of the identity can be verified by inspecting

$$W \longmapsto (I-P_1)AW, \quad \text{with } W\in\mathcal{W}, \tag{2}$$

where P_1 is a projection on $\mathbb{C}^{n\times n}$ onto $\mathcal{V}_1$ and I denotes the identity matrix (regarded as acting on $\mathbb{C}^{n\times n}$). If this linear map has a nontrivial nullspace with invertible elements W, equalities are attained and

$$A=V_1W^{-1} \tag{3}$$

holds with $V_1=P_1AW$.([3]) Such a factorization is of interest whenever the inverses of invertible elements of $\mathcal{W}$ admit a characterization.

With a characterization, the family of matrices representable as (3) yields a nonlinear structured subset of $\mathbb{C}^{n\times n}$. There appears a need for such geometries, for example, most recently in nonlinear dimensional reduction, where alternative structures to low rank approximations are constantly being seeked for better compression [30]. With an invertible $\mathcal{W}$, several classical examples were revisited in [21]. Otherwise the problem of characterization is challenging in general. Motivated by discretizations of partial differential equations, it has apparently been considered in a nontrivial case for the first time in [3]. Certainly, a lot of effort has been devoted to the task of describing the inverses of Toeplitz matrices; see [11], [13] and references therein. We show that $\mathcal{F}_k$, the set of matrices of rank at most k, which is probably the most encountered bilinear matrix family in practice, can be recovered (at the limit) with the inverses of invertible elements of a fairly large family of matrix subspaces. Related matrix structures are also looked at.

For the factorization (3) there arises a need to know how the nullspace of (2) possesses invertible elements. For the singular elements of a matrix subspace (of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$) with invertible elements, there is a natural concept of spectrum.([4]) Given by determinantal hypersurfaces in the complex projective space, this links matrix analysis and computations with classical topics in algebraic geometry.([5]) The weakest arising equivalence relation splits matrix subspaces into equivalence classes. Two of its stronger forms split the Grassmannian $\mathrm{Gr}_k(\mathbb{C}^{n\times n})$ consisting of k-dimensional subspaces of $\mathbb{C}^{n\times n}$ into equivalence classes. In an equivalence class

([3]) Finding the inverse of a nonsingular $A\in\mathbb{C}^{n\times n}$ can also be formulated as a matrix factorization problem once we set $\mathcal{W}=\mathbb{C}^{n\times n}$ and $\mathcal{V}_1=\mathbb{C}I$.

([4]) Naturality can be argued by the fact that it generalizes the concept of spectrum of a matrix.

([5]) For the terminology, in matrix analysis we follow [17] and in algebraic geometry [16].

the spectra coincide and the invertible elements can be regarded as having a closely related structure. From the viewpoint of matrix computations, to have a well-conditioned element in the factorization (3), a concept for conditioning of matrix subspaces is proposed. This conditioning is then related to a matrix factorization problem.

The paper is organized as follows. In Section 2 matrix subspaces with invertible elements are considered. In the invertible case the structure associated with the inverses can be characterized. The general nonlinear case is approached through examples. A fairly versatile structure is proposed whose inverses can be viewed as the product of three matrix subspaces. In Section 3 determinantal hypersurfaces are related with the singular elements of matrix subspaces. The spectrum of a matrix subspace is defined. Conditioning of matrix subspaces is addressed.

2. Matrix subspaces possessing invertible elements

In what follows we are concerned with (square) matrix subspaces possessing invertible elements.(6) Otherwise a matrix subspace is called *singular*. (Such subspaces are also of interest [10], [8].) If there exists an invertible element in a subspace $\mathcal{V}$ of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$), then the set of invertible elements is a dense and open subset of $\mathcal{V}$ [21, Theorem 2.2]. Such subspaces can be separated into two categories: by being either invertible, or merely possessing invertible elements.

2.1. Invertible matrix subspaces

Denote by $\mathrm{GL}(n,\mathbb{C})$ the group of invertible $n\times n$ complex-entried matrices and set

$$\mathrm{Inv}(\mathcal{V}) = \{V^{-1} : V \in \mathcal{V} \cap \mathrm{GL}(n,\mathbb{C})\}$$

for a matrix subspace $\mathcal{V}$ of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$) with invertible elements.(7) The so-called invertible matrix subspaces were introduced as an intermediate structure between a matrix subspace and a matrix subalgebra. They possess invertible elements and are confined to preserve the linear geometry under inversion in the following sense.

(6) In operator theory a closed subspace of $B(H)$, the algebra of all bounded operators on a Hilbert space H, is called an *operator space* [9]. It has grown to be an active area of research.

(7) Although we are primarily concerned with the geometry of $\mathrm{Inv}(\mathcal{V})$ for matrix subspaces, the structure is of interest for any set $\mathcal{V}$ of square matrices. An anonymous referee suggests looking at homogeneous sets, i.e., those $\mathcal{V}$ for which $t\mathcal{V}=\mathcal{V}$ for any $0\neq t\in\mathbb{C}$. He/she also suggests considering the adjugate operation instead of inversion. We are grateful for these interesting remarks.

Definition 2.1. Let $\mathcal{V}$ and $\mathcal{W}$ be two subspaces of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$). Then $\mathcal{W}$ is the *inverse* of $\mathcal{V}$ if

$$\mathrm{Inv}(\mathcal{V}) = \mathcal{W} \cap \mathrm{GL}(n,\mathbb{C}) \neq \varnothing.$$

Equivalently, the closure of $\mathrm{Inv}(\mathcal{V})$ is required to equal $\mathcal{W}$.

In the invertible case we use the notation $\mathcal{V}^{-1}=\mathcal{W}$. Certainly, if $\mathcal{V}$ is a subalgebra of $\mathbb{C}^{n\times n}$ possessing invertible elements, then always $\mathcal{V}^{-1}=\mathcal{V}$, i.e., $\mathcal{V}$ is closed under inversion.

Invertibility of a matrix $V\in\mathbb{C}^{n\times n}$ can be formulated equivalently in terms of invertibility of the respective matrix subspace $\mathcal{V}=\mathrm{span}\{V\}$ whose inverse in the invertible case is $\mathcal{V}^{-1}=\mathrm{span}\{V^{-1}\}$. Many classical concepts can be stated analogously, and more generally, in terms of matrix subspaces. (The spectrum is considered in Section 3.) For instance,

$$\mathcal{V} = \mathrm{span}\{I, V, ..., V^{k-1}\} \tag{4}$$

is invertible if and only if the degree of V satisfies $\deg(V)\leq k$.(8)

Example 2.2. Suppose $A\in\mathbb{C}^{m\times m}$, $B\in\mathbb{C}^{n\times n}$ and $C\in\mathbb{C}^{m\times n}$. Then solving the matrix equation $AX+YB=C$ is equivalent to the matrix factorization problem $\left[\begin{smallmatrix} A & C \\ 0 & B\end{smallmatrix}\right]=V_1V_2$ with V_1 and V_2 belonging to the matrix subspaces

$$\mathcal{V}_1 = \mathrm{span}\left\{\begin{pmatrix} A & YB \\ 0 & B\end{pmatrix} : Y\in\mathbb{C}^{m\times n}\right\} \quad \text{and} \quad \mathcal{V}_2 = \mathrm{span}\left\{\begin{pmatrix} I & X \\ 0 & I\end{pmatrix} : X\in\mathbb{C}^{n\times m}\right\},$$

respectively, see [12]. We have $\mathcal{V}_2^{-1}=\mathcal{V}_2$.(9)

A matrix subspace $\mathcal{V}$ of $\mathbb{C}^{n\times n}$ is said to be *polynomially closed* over $\mathbb{C}$ (or $\mathbb{R}$) if $p(V)\in\mathcal{V}$ for every $V\in\mathcal{V}$ and every polynomial p with complex (real) coefficients. In particular, a polynomially closed matrix subspace contains the scalars. A polynomially closed matrix subspace is invertible with $\mathcal{V}^{-1}=\mathcal{V}$ by the fact that the inverse of a nonsingular matrix V is a polynomial in V.

Example 2.3. The set of symmetric matrices is polynomially closed over $\mathbb{C}$. The set of Hermitian matrices is polynomially closed over $\mathbb{R}$. Hence they are both invertible matrix subspaces.

The converse does not hold.

---

(8) Matrix subspaces of this form are of great importance in matrix computations. Their elements have the rare property of being commutative.

(9) Linearizing the matrix factorization problem $A=V_1V_2$ at $(V_1,V_2)\in\mathcal{V}_1\times\mathcal{V}_2$ gives rise to the matrix equation $V_1X+YV_2=A-V_1V_2$.

Example 2.4. The matrix subspace $\mathcal{V}\subset\mathbb{C}^{n\times n}$, with $n=km$, spanned by

$$\left\{\begin{pmatrix} 0 & 0 & \dots & 0 & A_1 \\ I & 0 & \dots & 0 & A_2 \\ 0 & I & \dots & 0 & A_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & A_m \end{pmatrix} : A_j\in\mathbb{C}^{k\times k} \text{ for } j=1,\dots,m\right\}$$

is invertible with the inverse spanned by

$$\left\{\begin{pmatrix} B_1 & I & 0 & \dots & 0 \\ B_2 & 0 & I & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{m-1} & 0 & 0 & \dots & I \\ B_m & 0 & 0 & \dots & 0 \end{pmatrix} : B_j\in\mathbb{C}^{k\times k} \text{ for } j=1,\dots,m\right\}.$$

Clearly, $\mathcal{V}$ is not polynomially closed.

Next we demonstrate that being polynomially closed is always related to invertible matrix subspaces. For this, consider a subspace $\mathcal{V}$ of $\mathbb{C}^{n\times n}$ with invertible elements. The linearization of the inversion operation at an invertible matrix $\widehat{V}\in\mathcal{V}$ yields an approximation to the inverse of $\widehat{V}+V$ as

$$\widehat{V}^{-1}-\widehat{V}^{-1}V\widehat{V}^{-1}=\widehat{V}^{-1}(\widehat{V}-V)\widehat{V}^{-1} \quad \text{with } V\in\mathcal{V}, \tag{5}$$

hence giving rise to the matrix subspace $\widehat{V}^{-1}\mathcal{V}\widehat{V}^{-1}$, i.e., the tangent space of $\operatorname{Inv}(\mathcal{V})$ at $\widehat{V}^{-1}$. Since the map $V\mapsto V^{-1}$ from $\mathcal{V}\cap\mathrm{GL}(n,\mathbb{C})$ to $\mathbb{C}^{n\times n}$ is thus an injective immersion, we can conclude that $\operatorname{Inv}(\mathcal{V})$ is a submanifold of $\mathbb{C}^{n\times n}$. (For submanifold, see [27, p. 234].)

If $\mathcal{V}$ is invertible, then this linearization is independent of $\widehat{V}$ as follows.

Theorem 2.5. *If $\mathcal{V}$ is an invertible subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$), then*

$$\mathcal{V}^{-1}=V^{-1}\mathcal{V}V^{-1} \tag{6}$$

for any invertible $V\in\mathcal{V}$. Conversely, if $\mathcal{V}$ is a subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$) possessing invertible elements such that the right-hand side of (6) is independent of $V\in\mathcal{V}\cap\mathrm{GL}(n,\mathbb{C})$, then $\mathcal{V}$ is invertible.

Proof. Consider first the case of $\mathcal{V}$ being an invertible subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$. Suppose first that $\mathbb{C}I\subset\mathcal{V}$. We show that then $\mathcal{V}$ is polynomially closed. For this, look at the resolvent identity

$$(\mu I-V)^{-1}-(\lambda I-V)^{-1}=(\lambda-\mu)(\lambda I-V)^{-1}(\mu I-V)^{-1} \tag{7}$$

for any differing scalars λ and μ outside the spectrum of $V\in\mathcal{V}$. Since $\mathcal{V}^{-1}$ is a subspace, the left-hand side is in $\mathcal{V}^{-1}$. Therefore so is the right-hand side. Inverting both sides and subtracting an appropriate first-degree polynomial, we can conclude that V^2 is in $\mathcal{V}$. Therefore all the second-degree polynomials are in $\mathcal{V}$.

For higher powers, split a polynomial p as $p=p_1-p_2$. Then a generalization of the resolvent identity (7) (derived analogously) reads

$$(p_2(V)-q(V))^{-1}-(p_1(V)-q(V))^{-1}=p(V)(p_1(V)-q(V))^{-1}(p_2(V)-q(V))^{-1}$$

for any polynomial q for which the appearing inverses exist. With $p(V)=(\lambda-\mu)V$ and $q(V)=V^2$ we can conclude analogously that V^3 is in $\mathcal{V}$. Then with $p(V)=\lambda-\mu$ and $q(V)=V^2$ we can conclude that V^4 is in $\mathcal{V}$. Using this argument inductively, we can infer that all the polynomials in V are in $\mathcal{V}$. Thus $\mathcal{V}$ is polynomially closed.

Take now an invertible V. By the fact that V^{-1} is a polynomial in V, we can conclude that $V^{-1}\in\mathcal{V}$. Hence $\mathcal{V}^{-1}=\mathcal{V}$.

To prove finally that (6) holds, take any invertible $V\in\mathcal{V}$. Then $V^{-1}\mathcal{V}$ is an invertible subspace with the inverse $\mathcal{V}^{-1}V$. Since $\mathbb{C}I\subset V^{-1}\mathcal{V}$, we have $\mathcal{V}^{-1}V=V^{-1}\mathcal{V}$. Consequently, $\mathcal{V}^{-1}=V^{-1}\mathcal{V}V^{-1}$.

If $\mathbb{C}I\not\subset\mathcal{V}$, then take an invertible $V\in\mathcal{V}$ and consider $V^{-1}\mathcal{V}$. It is an invertible subspace with the inverse $\mathcal{V}^{-1}V$. Also $\mathbb{C}I\subset V^{-1}\mathcal{V}$. Therefore $\mathcal{V}^{-1}V=V^{-1}\mathcal{V}$, i.e., $\mathcal{V}^{-1}=V^{-1}\mathcal{V}V^{-1}$.

Consider now the case that $\mathcal{V}$ is an invertible subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{R}$. First suppose that $\mathbb{R}I\subset\mathcal{V}$. Then for any $V\in\mathcal{V}$, by analogous arguments, all the real polynomials in V are in $\mathcal{V}$. We have $p(V)=\sum_{j=0}^{2n}\alpha_jV^j=0$ for a monic polynomial p of the least possible degree with real coefficients. If V is invertible, then $\alpha_0\neq0$. Therefore $V(\sum_{j=1}^{2n}\alpha_jV^{j-1})=-\alpha_0I$, i.e., V^{-1} is a real polynomial in V and hence belongs to $\mathcal{V}$. Thus $\mathcal{V}^{-1}=\mathcal{V}$. The remainder of the proof proceeds as in the complex case.

The converse claim is immediate. □

Invertible matrix subspaces of dimension k in $\mathbb{C}^{n\times n}$ is hence the largest subset of the Grassmannian $\mathrm{Gr}_k(\mathbb{C}^{n\times n})$ with the property that the inversion operation is well-defined in $\mathrm{Gr}_k(\mathbb{C}^{n\times n})$.

Observe that for an invertible matrix subspace $\mathcal{V}$ there is no immediate way to improve the approximation (5) of the inverse of $\widehat{V}+V$ since it is already in $\mathcal{V}^{-1}$.

Corollary 2.6. *$\mathcal{V}$ is an invertible subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$) if and only if $\mathcal{V}=V\widehat{\mathcal{V}}$ with an invertible $V\in\mathbb{C}^{n\times n}$ and a matrix subspace $\widehat{\mathcal{V}}$ of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$) that is polynomially closed.*

In particular, if $\mathcal{V}$ contains the scalars, then $\mathcal{V}$ is invertible if and only $\mathcal{V}$ is polynomially closed. If not, then consider $\widehat{\mathcal{V}}=V^{-1}\mathcal{V}$ which contains the scalars for any invertible $V\in\mathcal{V}$.

Example 2.7. Let $\mathcal{V}\subset\mathbb{C}^{n\times n}$ be the subspace of Toeplitz matrices (over $\mathbb{C}$) with $n\geq 2$. We have $\mathbb{C}I\subset\mathcal{V}$. To see that $\mathcal{V}$ is not invertible, take $V\in\mathcal{V}$ with the entries t_j on the jth diagonal, for $-n+1\leq j\leq n-1$. Then the difference between the $(1,1)$ and $(2,2)$ entries of V^2 is $t_{n-1}t_{-n+1}-t_{-1}t_1$. Hence $V^2\notin\mathcal{V}$ generically. Consequently, $\mathcal{V}$ is not invertible.

Suppose a matrix subspace $\mathcal{V}$ possesses invertible elements without being invertible. For simplicity, assume $\mathcal{V}$ contains scalars. Then the smallest invertible matrix subspace containing $\mathcal{V}$ is constructible through forming polynomials in the elements of $\mathcal{V}$ and spanning them. The process is continued until a polynomially closed matrix subspace is obtained.

A matrix subspace $\mathcal{V}$ possessing invertible elements without being invertible has the property that the matrix subspace on the right in (6) depends on $V\in\mathcal{V}$.

Example 2.8. Analyzing the resolvent operator of $W\in\mathbb{C}^{n\times n}$ is equivalent to inspecting the submanifold $\operatorname{Inv}(\mathcal{V})$ for the matrix subspace $\mathcal{V}=\operatorname{span}\{I,W\}$. Then the right-hand side of (6) is

$$\operatorname{span}\{(\lambda I-W)^{-1},(\lambda I-W)^{-2}\}$$

and $\operatorname{span}\{I,W\}$ when V varies in $\mathcal{V}$. It remains fixed, i.e., we have an invertible subspace, if and only if $\deg(W)\leq 2$.

If $\mathcal{V}$ has invertible elements without being invertible, then $\operatorname{Inv}(\mathcal{V})$ is not readily characterizable by the fact that then the arising geometry is, by definition, nonlinear. Whenever a characterization (or partial) can be given, the matrix structure is intriguing.

2.2. Characterizing inverses and related matrix structures

When a subspace $\mathcal{V}\subset\mathbb{C}^{n\times n}$ possesses invertible elements without being invertible, the formula (6) is no longer applicable as such. Then giving a simple characterization of the submanifold $\operatorname{Inv}(\mathcal{V})$ is a nontrivial task. In what follows, we look at the problem of characterization in some special cases.

Example 2.9. For the matrix subspace of Hankel matrices, recall the elegant characterization according to which the inverses of its invertible elements coincide with the invertible Bézout matrices [11, Theorem 7.13].

The following can also be regarded as classical.

Example 2.10. The Cayley transform yields a characterization for the matrix subspace

$$\mathcal{V}=\{rI+iH : r\in\mathbb{R} \text{ and } H^*=H\in\mathbb{C}^{n\times n}\}$$

of $\mathbb{C}^{n\times n}$ over $\mathbb{R}$. Then the inverses of the invertible elements can be expressed as $(1/2r)(U+I)$ with $U=(rI-iH)(rI+iH)^{-1}$ unitary for $r\neq 0$. With $r=0$ we have the set of invertible skew-Hermitian matrices.

The Cayley transform and the respective characterization of the inverses exists also in other classical Lie algebras.

Occasionally a matrix subspace is naturally represented as the sum $\mathcal{V}_1+\mathcal{V}_2$ of two matrix subalgebras $\mathcal{V}_1$ and $\mathcal{V}_2$ of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$) possessing invertible elements. Associated with this, consider the sum

$$\mathcal{V}_1\mathcal{V}_2+\mathcal{V}_1\mathcal{V}_2. \tag{8}$$

Example 2.11. Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be the subalgebras of circulant and diagonal matrices of $\mathbb{C}^{n\times n}$. The matrix subspace $\mathcal{V}_1+\mathcal{V}_2$ appears in solving Schrödinger equations numerically [22]. The sum (8) arises in diffractive optics [21].

It can be more useful to view (8) alternatively as the product of three matrix subspaces as follows.

Proposition 2.12. *Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be two subalgebras of $\mathbb{C}^{n\times n}$ having invertible elements. Then the closures of (8) and*

$$\mathcal{V}_1(\mathcal{V}_1+\mathcal{V}_2)\mathcal{V}_2 \tag{9}$$

equal.

Proof. Since we are dealing with subalgebras, evidently any element of the space $\mathcal{V}_1(\mathcal{V}_1+\mathcal{V}_2)\mathcal{V}_2$ is an element of $\mathcal{V}_1\mathcal{V}_2+\mathcal{V}_1\mathcal{V}_2$.

Conversely, for any element $V_1V_2+\widehat{V}_1\widehat{V}_2\in\mathcal{V}_1\mathcal{V}_2+\mathcal{V}_1\mathcal{V}_2$, there are invertible elements $\widetilde{V}_1\in\mathcal{V}_1$ and $\widetilde{V}_2\in\mathcal{V}_2$ arbitrarily close to V_1 and $\widehat{V}_2$, respectively [21, Theorem 2.2]. Hence, the claim follows after factoring

$$\widetilde{V}_1(V_2\widetilde{V}_2^{-1}+\widetilde{V}_1^{-1}\widehat{V}_1)\widetilde{V}_2\in\mathcal{V}_1(\mathcal{V}_1+\mathcal{V}_2)\mathcal{V}_2. \quad \square$$

Consequently, the submanifold $\mathrm{Inv}(\mathcal{V}_1+\mathcal{V}_2)$ also plays a central role in characterizing the inverses of the sum $\mathcal{V}_1\mathcal{V}_2+\mathcal{V}_1\mathcal{V}_2$.

Example 2.13. Consider the set of circulant $\mathcal{V}_1$ and skew-circulant matrices $\mathcal{V}_2$, both subalgebras of $\mathbb{C}^{n\times n}$ with invertible elements. Then $\mathcal{V}_1+\mathcal{V}_2$ is the subspace of Toeplitz matrices. For its invertible elements, the inverse of a Toeplitz matrix has been shown to be in (8) [13]. Hence, with (9) the inverses of invertible elements of $\mathcal{V}_1\mathcal{V}_2+\mathcal{V}_1\mathcal{V}_2$ can be characterized. Also Example 2.9 can be used here since any Toeplitz matrix is the product of the backward identity and a Hankel matrix.([10])

The entries of (8) are of second degree while those of (9) are of third degree as polynomials in the entries of the factors. This is intriguing since an alternating iteration can be devised for approximating a matrix $A\in\mathbb{C}^{n\times n}$ with elements from (8) (and hence from (9)). For this, freeze the left multipliers and the right multipliers, alternatingly, and find the nearest element in the respective subspaces.

In view of this, consider the Kronecker product. For the Kronecker product of two matrix subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ we have

$$\mathrm{Inv}(\mathcal{V}_1\otimes\mathcal{V}_2)=\mathrm{Inv}(\mathcal{V}_1)\otimes\mathrm{Inv}(\mathcal{V}_2),$$

so that in the following special case we are dealing with matrix subspaces.

Proposition 2.14. *If $\mathcal{V}$ is an invertible matrix subspace, then $\mathcal{V}\otimes\mathbb{C}I$ and $\mathbb{C}I\otimes\mathcal{V}$ are invertible matrix subspaces.*

Suppose $kl=n$ and consider the subalgebras $\mathcal{V}_1=\mathbb{C}^{k\times k}\otimes\mathbb{C}I$ and $\mathcal{V}_2=\mathbb{C}I\otimes\mathbb{C}^{l\times l}$ of $\mathbb{C}^{n\times n}$. Then $\mathcal{V}_1+\mathcal{V}_2$ is the matrix subspace consisting of the Kronecker sums while $\mathcal{V}_1\mathcal{V}_2+\mathcal{V}_1\mathcal{V}_2$ consists of the sums of two Kronecker products. (For the Kronecker sum, see [18, Chapter 4.4].) As is well known, we have $e^{\mathcal{V}_1+\mathcal{V}_2}=\mathrm{GL}(k,\mathbb{C})\otimes\mathrm{GL}(l,\mathbb{C})$, i.e., the exponentials of $\mathcal{V}_1+\mathcal{V}_2$ are readily characterizable. We are not aware of a simple characterization of $\mathrm{Inv}(\mathcal{V}_1+\mathcal{V}_2)$, which is quite unsatisfactory.

Consider next $\mathcal{F}_k\subset\mathbb{C}^{n\times n}$, the set of matrices of rank at most k. Being a nonlinear matrix structure, let us look at it in terms of the product

$$(U,V)\longmapsto UV^*\quad\text{with } U,V\in\mathbb{C}^{n\times k}$$

and linearize this map at $(\widehat{U},\widehat{V})$. (Hence, $\mathcal{F}_k$ can also be regarded as the product of two matrix subspaces.) The constant term can be ignored (after making a change

([10]) The backward identity is otherwise a zero square matrix except that all its main antidiagonal entries equal one.

of variables $U = -\widehat{U} + \widetilde{U}$), giving us the matrix subspace

$$U\widehat{V}^* + \widehat{U}V^* \quad \text{with } U, V \in \mathbb{C}^{n\times k}, \tag{10}$$

which we denote by $\mathbb{C}^{n\times k}\widehat{V}^* + \widehat{U}\mathbb{C}^{k\times n}$. This is a subset of $\mathcal{F}_{2k}$. Possessing no invertible elements for $2k < n$, it appears natural to consider the matrix subspace consisting of the sum

$$\mathcal{W} = \mathcal{V} + \mathbb{C}^{n\times k}\widehat{V}^* + \widehat{U}\mathbb{C}^{k\times n} \tag{11}$$

for any fixed invertible subspace $\mathcal{V}$ of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$. This can be regarded as the tangent space of $\mathcal{V} + \mathcal{F}_k$ at a point.([11])

With the freedom to choose $\widehat{V}$, $\widehat{U}$ and $\mathcal{V}$, there is a fair amount of versatility in this family of matrix subspaces. Also, now $\mathrm{Inv}(\mathcal{W})$ admits a symmetric characterization in terms of the product of three matrix subspaces as follows.

Theorem 2.15. *Let $\mathcal{W}$ be defined as in (11). Then the closure of* $\mathrm{Inv}(\mathcal{W})$ *equals the closure of*

$$(\mathbb{C}I + \mathbb{C}^{n\times k}\widehat{V}^*)\mathcal{V}^{-1}(\widehat{U}\mathbb{C}^{k\times n} + \mathbb{C}I). \tag{12}$$

Proof. The matrix subspaces $\mathcal{V} + \mathbb{C}^{n\times k}\widehat{V}^*$ and $\mathcal{W}$ contain invertible elements since $\mathcal{V}$ does. By [21, Theorem 2.2], in each of these subspaces the set of invertible elements is open and dense.

Suppose $M + U\widehat{V}^* + \widehat{U}V^* \in \mathcal{W}$ is invertible. We may assume, after an arbitrary small perturbation, if necessary, that M and $M + U\widehat{V}^*$ are invertible as well. Let $N = I + \widehat{U}V^*(M + U\widehat{V}^*)^{-1}$. Then the inverse can be expanded as

$$\begin{aligned}(M + U\widehat{V}^* + \widehat{U}V^*)^{-1} &= (N(M + U\widehat{V}^*))^{-1}\\ &= (M + U\widehat{V}^*)^{-1}N^{-1} = (I + M^{-1}U\widehat{V}^*)^{-1}M^{-1}N^{-1}.\end{aligned}$$

With this, recall that, for $X, Y \in \mathbb{C}^{n\times k}$, the matrix $I + XY^*$ is invertible if and only if $(I + XY^*)^{-1} = I - X(I + Y^*X)^{-1}Y^*$. Therefore the first and second factors in the expansion can be written as $I - \widetilde{U}\widehat{V}^*$ and $I - \widehat{U}\widetilde{V}^*$ with $\widetilde{U}, \widetilde{V} \in \mathbb{C}^{n\times k}$. This proves that the closure of $\mathrm{Inv}(\mathcal{W})$ belongs to the closure of (12).

The sets $(\mathbb{C}I + \mathbb{C}^{n\times k}\widehat{V}^*)\mathcal{V}^{-1}(\mathbb{C}I + \widehat{U}\mathbb{C}^{k\times n})$ and $(I + \mathbb{C}^{n\times k}\widehat{V}^*)\mathcal{V}^{-1}(I + \widehat{U}\mathbb{C}^{k\times n})$ have the same closures since $\mathcal{V}$ is a subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$. Therefore, reversing the steps yields the converse inclusion. □

With the following choices $\mathcal{F}_k$ is recovered with the product (12).

([11]) $\mathcal{V} + \mathcal{F}_k$ can be regarded as the set of “small rank perturbations” of the matrix subspace $\mathcal{V}$. Such structures have been studied, e.g., in [19].

Corollary 2.16. *Suppose there exists* $M\in\mathcal{V}^{-1}$ *such that* $\widehat{V}^*M\widehat{U}$ *is invertible. Then the set* (12) *contains* $\mathcal{F}_k$.

Proof. For any $\alpha,\beta\in\mathbb{C}$ and $M\in\mathcal{V}^{-1}$ we have

$$(13)\quad (\alpha I+\mathbb{C}^{n\times k}\widehat{V}^*)M(\beta I+\widehat{U}\mathbb{C}^{k\times n}) = \alpha\beta M+\alpha M\widehat{U}\mathbb{C}^{k\times n}+\beta\mathbb{C}^{n\times k}\widehat{V}^*M+\mathbb{C}^{n\times k}\widehat{V}^*M\widehat{U}\mathbb{C}^{k\times n}.$$

Setting $\alpha=\beta=0$ yields $\mathcal{F}_k$, whenever $\widehat{V}^*M\widehat{U}$ is invertible. □

For the simplest option (but a very important one in view of applications), choose $\mathcal{V}=\mathbb{C}I$ and, let us say, $\widehat{U}=\widehat{V}=\binom{I}{0}\in\mathbb{C}^{n\times k}$ with the identity matrix I of size $k\times k$. Then the matrix subspace $\mathcal{W}$ consumes the same amount of storage as $\mathcal{F}_k$. For a fixed $A\in\mathbb{C}^{n\times n}$, this can be viewed as providing an alternative to the structure appearing in the approximation problem

$$(14)\qquad \min_{\substack{\lambda\in\mathbb{C}\\ U,V\in\mathbb{C}^{n\times k}}} \|A-\lambda I-UV^*\|.$$

We are not aware of a formula or a simple way of solving this with the canonical forms.([12]) As opposed to this, approximations in the residual sense with $\mathcal{W}$ can be found. For this, take the linear map (2), choose $\mathcal{V}_1=\mathbb{C}I$ and consider

$$\min_{\substack{W\in\mathcal{W}\\ \|W\|_F=1}} \|(I-P_1)AW\|_F.$$

(See also [21, equation (2.6)].) This problem can be solved by invoking the singular value decomposition of (2). Since this yields $AW\approx I$, the actual approximation of A is in $\operatorname{Inv}(\mathcal{W})$ whose closure is, according to (12), a subset of $\mathbb{C}I+\mathcal{F}_{2k}$.

Certainly, the inverses of invertible elements do not necessarily need a particular characterization even though solving (1) approximately and subsequently finding the inverse is of interest. In squaring and scaling for the matrix exponential this arises in the form of a rational approximation problem

$$(15)\qquad A=e^V\approx p(V)q(V)^{-1}$$

for the exponential of $V\in\mathbb{C}^{n\times n}$, with polynomials p of degree k and q of degree l at most. These are then the respective matrix subspaces $\mathcal{V}_1$ and $\mathcal{W}$ defined as in (4).

For another illustration, with iterative methods for solving very large linear systems, there arises the problem of approximating

$$(16)\qquad AWV_1^{-1}\approx I$$

([12]) For matrices from $\mathbb{C}I+\mathcal{F}_k$, the exponential can be computed fast [20].

for an invertible $A \in \mathbb{C}^{n\times n}$ with invertible preconditioning matrices $W \in \mathcal{W}$ and $V_1 \in \mathcal{V}_1$. Now the subspace $\mathcal{V}_1$ is required to consist of matrices for which it is possible to solve linear systems fast, whenever V_1 is invertible [7].

3. The spectrum and conditioning of matrix subspaces

For the factorization (3) there arises a need to understand how the nullspace of (2) possesses invertible elements. Consequently, in what follows, we are concerned with the singular elements of a matrix subspace $\mathcal{W}$ of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ possessing invertible elements. The real case $\mathcal{W} \subset \mathbb{R}^{n\times n}$ is not considered here.(13) Belonging to the realm of real algebraic geometry, it leads to very challenging problems requiring different techniques. (For an illustration of this, see Example 3.3 below.)

3.1. The spectrum of a matrix subspace

For the singular elements, suppose $W_1, ..., W_k$ is a basis of the matrix subspace $\mathcal{W} \subset \mathbb{C}^{n\times n}$ and set

$$p(z_1, ..., z_k) = \det(z_1 W_1 + ... + z_k W_k) \tag{17}$$

for $z = (z_1, ..., z_k) \in \mathbb{C}^k$. This is a homogeneous polynomial of degree n. Therefore

$$V(p) = \{z \in \mathbb{C}^k : p(z_1, ..., z_k) = 0\} \tag{18}$$

determines a hypersurface in the $(k-1)$-dimensional projective space $\mathbb{P}^{k-1}(\mathbb{C})$. (Hence we assume $k \geq 2$.) For obvious reasons, such hypersurfaces are called *determinantal*.

Definition 3.1. The hypersurface (18) is called the *spectrum* of $\mathcal{W}$ in the basis $W_1, ..., W_k$.

The spectrum is always nonempty and continuous. For these claims, see Appendix A for the metric used and the proof.

Although $V(p)$ depends on the basis used, any other such a hypersurface is obtained through a change of variables as $XV(p)$ with an invertible $X \in \mathbb{C}^{k\times k}$. Such hypersurfaces are said to be projectively equivalent. It is readily seen that they yield all the possible spectra of $\mathcal{W}$.

(13) The real case is of equal importance in view of real matrix factorization problems.

Example 3.2. Consider the matrix subspace of Example 2.4 with $k=1$, i.e., the span of companion matrices. Let W_1 be the nilpotent forward shift and W_j otherwise be zero except that the $(j-1,n)$ entry equals one, for $j=2,...,n+1$. Then $p(z_1,...,z_{n+1})=(-1)^{n-1}z_1^{n-1}z_2$ yielding readily $V(p)$.

The concept is also related to the following "property P" arising in scattering theory.

Example 3.3. Motivated by hyperbolic partial differential equations, in [1] it was studied when the spectrum of $\mathcal{W}$ in a certain given basis $W_1,...,W_k$ does not intersect $\mathbb{P}^{k-1}(\mathbb{R})$. (More precisely, how large k can be such that this is still possible.(14)) When $\mathcal{W}\subset\mathbb{R}^{n\times n}$, this illustrates how problems in real algebraic geometry are different, i.e., all the nonzero elements of a matrix subspace can be invertible.

Projective equivalence splits matrix subspaces into equivalence classes as follows.

Definition 3.4. Matrix subspaces $\mathcal{W}$ and $\mathcal{V}$ are *projectively equivalent* if their spectra are projectively equivalent.(15)

This is seemingly the weakest equivalence relation of interest. A stronger form consists of requiring the polynomials (17) to coincide in some bases of $\mathcal{W}$ and $\mathcal{V}$. Then $\mathcal{W}$ and $\mathcal{V}$ are said to be *strongly projectively equivalent*. This forces, for instance, $\mathcal{W}$ and $\mathcal{V}$ to be in the same $\mathbb{C}^{n\times n}$. An even stronger equivalence relation requires $\mathcal{W}=X\mathcal{V}Y$ to hold for some invertible matrices $X,Y\in\mathbb{C}^{n\times n}$. Then the matrix subspaces $\mathcal{W}$ and $\mathcal{V}$ are said to be *equivalent*.(16)

With these two stronger equivalence relations, it is natural to work with the Grassmannian $\mathrm{Gr}_k(\mathbb{C}^{n\times n})$ consisting of k-dimensional subspaces of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$, for $1\le k\le n^2$.

Example 3.5. The set of Toeplitz and the set of Hankel matrices are equivalent. For this, take X to be the backward identity and $Y=I$.

Theorem 2.5 has the following corollary.

(14) Zeros of real homogeneous polynomials in a large number of variables have received attention recently; see [2] and references therein.

(15) Although they must be of the same dimension, we do not require $\mathcal{W}$ and $\mathcal{V}$ to be in the same $\mathbb{C}^{n\times n}$.

(16) This complies with the terminology used in connection with the generalized eigenvalue problem.

Theorem 3.6. *An invertible matrix subspace is equivalent to its inverse.*

For one more instructive example, set $\mathcal{W}^T=\{W^T : W\in\mathcal{W}\}$. By elementary properties of the determinant, any matrix subspace $\mathcal{W}$ is strongly projectively equivalent to $\mathcal{W}^T$.

Certainly, to actually recover the singular elements from a spectrum of $\mathcal{W}$, the associated basis must be known. This assumption is implicitly made with the spectrum of a matrix. Then a fixed basis is always employed, guaranteeing the uniqueness of the concept.

Example 3.7. The spectrum of a matrix $W\in\mathbb{C}^{n\times n}$ is related to the matrix subspace $\mathcal{W}=\operatorname{span}\{I,W\}$. Typically, the elements of a matrix subspace are regarded as indistinguishable whereas with this particular $\mathcal{W}$, the identity matrix is assigned a special role. This means that one is concerned with the determinant of $\lambda I-W$ as opposed to (17). Hence the basis used with the spectrum is always $W_1=I$ and $W_2=-W$. Then

$$\det(\lambda I-W)=\prod_{j=1}^{n}(\lambda-\lambda_j) \quad\Longleftrightarrow\quad p(z_1,z_2)=\prod_{j=1}^{n}(z_1-\lambda_j z_2),$$

where λ_j denote the eigenvalues of W.

Given a matrix subspace, it is an intriguing question, which type of spectra can arise. In algebraic geometry one is concerned with the related question, whether a given hypersurface is determinantal or not, typically in the subspace of symmetric matrices; see [4], [26] and references therein.

For the two-dimensional case, let W_1 and W_2 be a basis of $\mathcal{W}$ with W_1 being invertible. Then $\mathcal{W}$ is equivalent to $\operatorname{span}\{I,W_1^{-1}W_2\}$ and therefore it suffices to consider the following case.

Theorem 3.8. *For two matrices V and W,* $\operatorname{span}\{I,V\}$ *and* $\operatorname{span}\{I,W\}$ *are projectively equivalent if and only if the eigenvalues* (*in some order*, *not counting multiplicities*) *of V and W are related through a linear fractional transformation.*

Proof. Assume first that V and W are invertible. Denote by $\lambda_j(V)$ (resp. $\lambda_j(W)$) the distinct eigenvalues of V (resp. W). Then $\operatorname{span}\{I,V\}$ and $\operatorname{span}\{I,W\}$ are projectively equivalent if and only if

$$X\begin{pmatrix}\lambda_j(V)\\ 1\end{pmatrix}=Y\begin{pmatrix}\lambda_j(W)\\ 1\end{pmatrix} \tag{19}$$

for some invertible $X, Y \in \mathbb{C}^{2\times 2}$ and for the distinct eigenvalues in some order. Writing $Y^{-1}X = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, this gives

$$\lambda_j(W) = \frac{a\lambda_j(V)+b}{c\lambda_j(V)+d}$$

after dividing the equations.

In the singular case, proceed analogously with translations of V and W which are invertible. □

Corollary 3.9. *Suppose $V, W \in \mathbb{C}^{n\times n}$. Then* span$\{I, V\}$ *and* span$\{I, W\}$ *are strongly projectively equivalent if and only if the eigenvalues (in some order) of V and W are related through a linear fractional transformation such that the corresponding eigenvalues have the same algebraic multiplicities.*

The matrix subspaces $\mathcal{W}$ and $X\mathcal{W}Y$ of dimension k are equivalent for any invertible $X, Y \in \mathbb{C}^{n\times n}$. In particular, if there exist invertible $X, Y \in \mathbb{C}^{n\times n}$ such that XW_jY are upper triangular for $j=1, \ldots, k$, then the spectrum of $\mathcal{W}$ is a union of at most n linear varieties, in any basis. For $k=2$ this is always so. (Recall also that any homogeneous polynomial in two complex variables factors completely into linear homogeneous polynomials.) For $k \geq 3$ this is no longer true in general.

By employing the Jordan canonical form, this yields us the following proposition.

Proposition 3.10. *For any $V \in \mathbb{C}^{n\times n}$ the spectrum of the polynomial matrix subspace (4) in any basis is a union of at most n linear varieties.*

3.2. Conditioning of matrix subspaces

To compute a factorization (3) in practice, knowing the spectrum of a matrix subspace $\mathcal{W}$, or that of the nullspace of (2), in some basis is not quite sufficient. For the numerical stability of matrix computations, there should exist elements whose inverses are reliably computable. Recall that the condition number of a matrix is the ratio of its largest and smallest singular values. The best conditioned matrices are unitary. Hence the question arises, how near to unitary matrices can we get by scaling. (Scaling is a standard operation before executing the Gaussian elimination for the inverse [14].)

Although a very interesting problem, finding a best conditioned element of $\mathcal{W}$ seems very challenging as soon as $\dim(\mathcal{W}) > 1$. For a more tractable tool, scale the determinant function by setting

$$h(W)=\frac{|\det(W)|}{\prod_{j=1}^{n}\|w_j\|}, \tag{20}$$

where w_j denotes the jth column of $W\in\mathbb{C}^{n\times n}$.(17) It can be instructive to alternatively write $h(W)=|\det(WD)|$ with $D=\operatorname{diag}(1/\|w_1\|,\dots,1/\|w_n\|)$. By Hadamard's inequality [17, Corollary 7.8.2], $0\le h(W)\le 1$ with the latter equality holding if and only if W has orthogonal columns, i.e., there exists a factorization

$$W=UD \tag{21}$$

of W with a unitary U and an invertible diagonal matrix D.

Proposition 3.11. *Suppose $W\in\mathbb{C}^{n\times n}$ has the singular value decomposition $W=U\Sigma V^*$ with V having entries of equal modulus. Then $h(W)=n^{n/2}\prod_{j=1}^{n}\sigma_j/(\sum_{j=1}^{n}\sigma_j^2)^{n/2}$.*

Proof. For the numerator, we have $|\det(W)|=\prod_{j=1}^{n}\sigma_j$. For the denominator, $\|w_j\|=\|(\Sigma V^*)_j\|=(1/\sqrt{n})(\sum_{j=1}^{n}\sigma_j^2)^{1/2}$. □

Example 3.12. Suppose $W\in\mathbb{C}^{n\times n}$ is circulant. Then the assumptions of Proposition 3.11 are satisfied and $h(W)=n^{n/2}\prod_{j=1}^{n}|\lambda_j|/(\sum_{j=1}^{n}|\lambda_j|^2)^{n/2}$, where λ_j, for $j=1,\dots,n$, denote the eigenvalues of W.

If $h(W)$ is not very small, then there are unitary matrices nearby as follows.

Theorem 3.13. *An invertible $W\in\mathbb{C}^{n\times n}$ can be scaled from the right with a diagonal matrix D such that*

$$WD=Q+\Delta$$

with Q unitary and $\|\Delta\|_F\le\sqrt{1/h(W)^2-1}$.

Proof. Let $W=QR$ be the QR-factorization of W with the diagonal entries of R satisfying $r_{jj}>0$. Inserting this into (20) yields $\prod_{j=1}^{n}r_{jj}=h(W)\prod_{j=1}^{n}\|r_j\|$ by the properties of the determinant and by the fact that $\|w_j\|=\|r_j\|$. Since $h(W)$ is invariant under scalings from the right, take $D=\operatorname{diag}(1/r_{11},\dots,1/r_{nn})$ to have

$$1=h(W)\prod_{j=1}^{n}\|r_j/r_{jj}\|. \tag{22}$$

(17) Here h refers to Hadamard by the fact that the so-called Hadamard number of W is defined as $1/h(W)$ [6].

Since the jth entry of r_j/r_{jj} is 1, write $\|r_j/r_{jj}\|^2=1+x_j$ with $x_j\geq 0$. Then from the identity (22) we have

$$\frac{1}{h(W)^2}=\prod_{j=1}^{n}(1+x_j)\geq 1+\sum_{j=1}^{n}x_j$$

so that $\sum_{j=1}^{n}x_j\leq 1/h(W)^2-1$. Therefore $WD=QRD=Q(I+\tilde{\Delta})$ with the norm satisfying $\|\tilde{\Delta}\|_F\leq\sqrt{1/h(W)^2-1}$ proving the claim by the unitary invariance of the Frobenius norm. □

Recall that in the operator norm

$$\min_{U\text{ unitary}}\|WD-U\|=\max_{1\leq j\leq n}|\sigma_j(WD)-1|$$

(see [17]). Hence, with the diagonal matrix D used in the proof, the condition number satisfies

$$\varkappa(WD)\leq\frac{1+\sqrt{1/h(W)^2-1}}{1-\sqrt{1/h(W)^2-1}}.$$

When the diagonal entries of D are chosen to equal the reciprocals of the norms of the columns, then $\varkappa(WD)\leq 2/h(W)$ [15].

Consider a matrix subspace $\mathcal{W}$. Fix a basis $W_1,...,W_k$ of $\mathcal{W}$. Then, in the complement of $V(p)$, the map

$$(z_1,...,z_k)\longmapsto h\left(\sum_{j=1}^{k}z_jW\right)^2 \tag{23}$$

admits computation of its gradient([18]) in a closed form. Consequently, finding now the extreme points (with a descent method) is a more tractable problem as opposed to finding a best conditioned element from $\mathcal{W}$.

Definition 3.14. The scaled condition number of a matrix subspace $\mathcal{W}\subset\mathbb{C}^{n\times n}$ is the square root of the supremum of the map (23).

This quantity obviously does not depend on the basis used.

([18]) Here $\mathbb{C}^k$ is identified with $\mathbb{R}^{2k}$.

Example 3.15. Let $\mathcal{W}\subset\mathbb{C}^{n\times n}$ be the subspace of upper triangular matrices. For $j\leq k$, take $W_{jk}=e_je_k^*$ to be its basis, where e_j denotes the jth standard basis vector of $\mathbb{C}^n$. Then

$$h\Big(\sum_{k=1}^{n}\sum_{j=1}^{k} z_{jk}W_{jk}\Big)^2=\frac{\prod_{j=1}^{n}|z_{jj}|^2}{\prod_{k=1}^{n}\sum_{j=1}^{k}|z_{jk}|^2}$$

whose critical points are readily seen to be located at the subspace of diagonal matrices. There the function attains the constant value 1 (whenever invertible) which cannot be improved. All this was undoubtedly evident since there are diagonal unitary matrices in $\mathcal{W}$.

3.3. Matrix factorization problem and conditioning of matrix subspaces

To formulate the matrix factorization problem (21) in more general terms, let $\mathcal{V}_1$ denote the set of unitary matrices and suppose that $\mathcal{V}_2$ is an invertible matrix subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ (or $\mathbb{R}$). Denote the inverse of $\mathcal{V}_2$ by $\mathcal{W}$. Then, for a given nonsingular matrix $M\in\mathbb{C}^{n\times n}$, the question of whether there exists a factorization

$$M=V_1V_2 \quad \text{with } V_1\in\mathcal{V}_1 \text{ and } V_2\in\mathcal{V}_2 \tag{24}$$

turns into a question concerning the conditioning of a matrix subspace.[19]

Namely, the existence of a factorization (24) is equivalent to having $M^*M=V_2^*V_2$. For this latter identity, denote the positive definite matrix M^*M by A and consider the linear map (2) with the matrix subspaces $\mathcal{V}_1=\mathcal{W}^*$ and $\mathcal{W}=\mathcal{V}_2$. Assume that its nullspace $\mathcal{N}$ possesses invertible elements. Then $A=V_2^*V_2$ if and only if $(AV)^*=V^{-1}$ for an invertible element V in the nullspace. This yields us the following characterization.

Proposition 3.16. *Let $A=XDX^*$ be positive definite with a unitary X and a diagonal matrix D. Then, for an invertible V there holds $(AV)^*=V^{-1}$ if and only if V^*X can be factored as* (21).

Proof. We have

$$(AV)^*=V^{-1} \quad\Longleftrightarrow\quad V^*AV=I \quad\Longleftrightarrow\quad (D^{1/2}N)^*(D^{1/2}N)=I$$

with $N=X^*V$. The last identity holds if and only if $D^{1/2}N$ is unitary. □

Consequently, the question of whether M can be factored as (24) converts into the question of whether the matrix subspace $\mathcal{N}^*X=\{N^*X : N\in\mathcal{N}\}$ possesses perfectly scaled elements measured in terms of (20).

[19] Besides (24), also the QR-factorization is a problem of this type.

4. Conclusions

A square matrix subspace with invertible elements is either invertible or merely possesses invertible elements. The former case can be regarded as well-understood while for the latter there does not seem to be a general way to characterize the inverses. For the singular elements of a matrix subspace there exists a concept for the spectrum. This allows for classifying matrix subspaces in terms of projective equivalence. Scaling the determinant function yields a criterion for finding stably invertible elements from a matrix subspace.

Appendix A

To prove that the spectrum is always nonempty, we need to show that (18) contains nonzero points $(z_1,...,z_k)$. The claim is clear if some W_j is singular. So let us assume that all $W_1,...,W_k$ are invertible. Then it suffices to put $z_2=1$ and $z_3=...=z_k=0$ and look at the univariate polynomial $q(z_1)=\det(z_1W_1W_2^{-1}+I)$ which has zeros by the fundamental theorem of algebra. Clearly, all these zeros have a strictly positive modulus.

Although seldom employed in matrix analysis, the following concepts are standard in algebraic geometry.

On $\mathbb{P}^{k-1}(\mathbb{C})$ the Kähler metric is used. For the Kähler metric, see [23, p. 247].

On any matrix subspace $\mathcal{W}\subset\mathbb{C}^{n\times n}$ the inner product

$$(W_1,W_2)=\operatorname{tr}(W_2^*W_1)$$

is used, for $W_1,W_2\in\mathcal{W}$.

With this inner product, consider the Grassmannian $\mathrm{Gr}_k(\mathbb{C}^{n\times n})$. Suppose $\mathcal{W},\mathcal{V}\in\mathrm{Gr}_k(\mathbb{C}^{n\times n})$ and denote by $P_{\mathcal{V}}$ and $P_{\mathcal{W}}$ the orthogonal projectors on $\mathbb{C}^{n\times n}$ onto $\mathcal{W}$ and $\mathcal{V}$. Then also

$$d(\mathcal{W},\mathcal{V})=\|P_{\mathcal{W}}-P_{\mathcal{V}}\|_2$$

yields a metric on $\mathrm{Gr}_k(\mathbb{C}^{n\times n})$, where $\|\cdot\|_2$ denotes the operator norm.

In the Grassmannian $\mathrm{Gr}_k(\mathbb{C}^{n\times n})$, suppose a sequence $\mathcal{V}_j$ converges to $\mathcal{W}$. Then a sequence $\{V_{j,1},...,V_{j,k}\}$ of bases of $\mathcal{V}_j$ is said to converge to a basis $W_1,...,W_k$ of $\mathcal{W}$ if

$$\lim_{l\to\infty}V_{j,l}=W_j$$

for $l=1,...,k$.

Proposition. *Suppose a sequence* $\{V_{j,1},...,V_{j,k}\}$ *of bases of* $\mathcal{V}_j$ *converges to a basis* $W_1,...,W_k$ *of* $\mathcal{W}$. *Then the sequence* $V(p_j)$ *of the respective spectra converges to* $V(p)$.

Proof. Assume that $W=\sum_{l=1}^{k} z_l W_l$ and $V_j=\sum_{l=1}^{k} \mu_{j,l} V_{j,l}$. Then $\lim_{j\to\infty} \mu_{j,l}=z_l$ for every $l=1,...,k$ if and only if $\lim_{j\to\infty} V_j=W$.

If W is invertible, then so is $V_j=W(I-W^{-1}(W-V_j))$ for $\|W^{-1}(W-V_j)\|<1$. Hence a sequence of singular elements cannot converge to W and therefore $V(p_j)$ do not converge to any points in the complement of $V(p)$.

Suppose next that W is singular and there does not exist a subsequence of V_j of singular elements. Necessarily, though, $\lim_{j\to\infty} \|V_j^{-1}\|_2=\infty$. Hence, V_j+E_j is singular with E_j having the Frobenius norm $1/\|V_j^{-1}\|_2$. Expand $E_j=\sum_{l=1}^{k} \varepsilon_{j,l} V_{j,l}$. Since each $V_{j,l}$ converges as $j\to\infty$, necessarily $\lim_{j\to\infty} \varepsilon_{j,l}=0$. □

References

1. Adams, J., Lax, P. and Phillips, R., On matrices whose real linear combinations are nonsingular, *Proc. Amer. Math. Soc.* **16** (1965), 318–322.
2. Aron, R. M. and Hájek, P., Zero sets of polynomials in several variables, *Arch. Math.* (*Basel*) **86** (2006), 561–568.
3. Asplund, E., Inverses of matrices a_{ij} which satisfy $a_{ij}=0$ for $j>i+p$, *Math. Scand.* **7** (1959), 57–60.
4. Beauville, A., Determinantal hypersurfaces, *Michigan Math. J.* **48** (2000), 39–64.
5. Bhatia, R., Matrix factorizations and their perturbations, *Linear Algebra Appl.* **197**, **198** (1994), 245–276.
6. Birkhoff, G., Two Hadamard numbers for matrices, *Comm. ACM* **18** (1975), 25–29.
7. Byckling, M. and Huhtanen, M., Approximate factoring of the inverse, *Preprint*, 2008.
8. Causin, A. and Pirola, G. P., A note on spaces of symmetric matrices, *Linear Algebra Appl.* **426** (2007), 533–539.
9. Effros, E. G. and Ruan, Z.-J., *Operator Spaces*, Oxford University Press, New York, 2000.
10. Falikman, D., Friedland, S. and Loewy, R., On spaces of matrices containing a nonzero matrix of bounded rank, *Pacific J. Math.* **207** (2002), 157–176.
11. Fiedler, M., *Special Matrices and Their Applications in Numerical Mathematics*, Martinus Nijhoff Publishers, Dordrecht, 1986.
12. Flanders, H. and Wimmer, H., On the matrix equations $AX-XB=C$ and $AX-YB=C$, *SIAM J. Appl. Math.* **32** (1977), 707–710.
13. Gohberg, I. and Olshevsky, V., Circulants, displacements and decompositions of matrices, *Integral Equations Operator Theory* **15** (1992), 730–743.
14. Golub, G. H. and Van Loan, C. F., *Matrix Computations*, 3rd ed., The John Hopkins University Press, Baltimore, 1996.

15. Guggenheimer, H., Edelman, A. and Johnson, C. R., A simple estimate of the condition number of a linear system, *College Math. J.* **26** (1995), 2–5.
16. Harris, J., *Algebraic Geometry*, Springer, New York, 1992.
17. Horn, R. A. and Johnson, C. R., *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1987.
18. Horn, R. A. and Johnson, C. R., *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1991.
19. Huhtanen, M., A matrix nearness problem related to iterative methods, *SIAM J. Numer. Anal.* **39** (2001), 407–422.
20. Huhtanen, M., Averaging operators for exponential splittings, *Numer. Math.* **106** (2007), 511–528.
21. Huhtanen, M., Factoring matrices into the product of two matrices, *BIT* **47** (2007), 793–808.
22. Jahnke, T. and Lubich, C., Error bounds for exponential operator splittings, *BIT* **40** (2000), 735–744.
23. Kostrikin, A. I. and Manin, Yu., *Linear Algebra and Geometry*, Gordon and Breach Science Publ., New York, 1989.
24. Laffey, T. J., Conjugacy and factorization results on matrix groups, in *Functional Analysis and Operator Theory* (*Warsaw, 1992*), Banach Center Publications **30**, pp. 203–221, Polish Academy of Sciences, Warsaw, 1994.
25. Lax, P., On the factorization of matrix valued functions, *Comm. Pure Appl. Math.* **24** (1976), 683–688.
26. Piontkowski, J., Linear symmetric determinantal hypersurfaces, *Michigan Math. J.* **54** (2006), 117–146.
27. Postnikov, M., *Lectures in Geometry, Smooth Manifolds*, Nauka, Moscow, 1987 (Russian). English transl.: Mir, Moscow, 1989.
28. Van Loan, C. and Pitsianis, N., Approximation with Kronecker products, in *Linear Algebra for Large Scale and Real-Time Applications* (*Leuven, 1992*), pp. 293–314, Kluwer, Dordrecht, 1993.
29. Wu, P. Y., The operator factorization problems, *Linear Algebra Appl.* **117** (1989), 35–63.
30. Ye, Q., Zha, H. and Li, R.-C., Analysis of an alignment algorithm for nonlinear dimensionality reduction, *BIT* **47** (2007), 873–885.

Marko Huhtanen
Institute of Mathematics
Helsinki University of Technology
P.O. Box 1100
FI-02015 Helsinki University of Technology
Finland
Marko.Huhtanen@hut.fi

Received April 30, 2008
published online June 3, 2009