Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 421 (2007) 163–170
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Author's personal copy

On the angular metrics between linear subspaces ☆

Yanxia Zhang [a], Li Qiu [b,*]

[a] Department of Mathematics, Beijing Institute of Technology, Beijing 100081, China
[b] Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong, China

Received 1 November 2005; accepted 8 August 2006
Available online 12 October 2006
Submitted by R. Bhatia

Abstract

It has recently been shown that the symmetric gauge functions on the canonical (principal) angles give a family of unitarily invariant metrics between linear subspaces of the same dimension. In this short paper, we extend such metrics to subspaces of possibly different dimensions. This extension is necessary in addressing some perturbation analysis problems involving subspaces with different dimensions. Examples of such perturbation analysis problems are also studied in this paper using the extended metrics.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Subspaces; Unitarily invariant metric; Canonical angles; Perturbation; Nullity; Deficiency

1. Introduction

Let $\mathscr{X}$ and $\mathscr{Y}$ be m- and l-dimensional subspaces of $\mathbb{C}^n$, respectively. The *canonical angles* between them are defined to be

$$\theta_i(\mathscr{X}, \mathscr{Y}) = \arccos \sigma_{\min\{m,l\}-i+1}(X^H Y), \quad i = 1, 2, \ldots, \min\{m, l\}, \tag{1}$$

☆ This work was supported by the Basic Research Funds of Beijing Institute of Technology under grant BIT-UBF-200507A4215 and the Hong Kong Research Grants Council under grant HKUST6171/02E. Part of it was done during the second author's visit to the Shanghai Key Laboratory of Contemporary Applied Mathematics, Fudan University, Shanghai, China.
* Corresponding author. Tel.: +852 2358 7067; fax: +852 2358 1485.
E-mail addresses: yxzhang@bit.edu.cn (Y. Zhang), eeqiu@ust.hk (L. Qiu).

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.08.015

where X and Y are matrices whose columns form orthonormal bases of $\mathscr{X}$ and $\mathscr{Y}$, and $\sigma_i(X^HY)$, $i = 1, 2, \ldots, \min\{m, l\}$, are decreasingly ordered singular values of X^HY. In the following, we denote the $\min\{m, l\}$-tuple of canonical angles between $\mathscr{X}$ and $\mathscr{Y}$ by

$$\theta(\mathscr{X}, \mathscr{Y}) := (\theta_1(\mathscr{X}, \mathscr{Y}), \ldots, \theta_{\min\{m,l\}}(\mathscr{X}, \mathscr{Y})).$$

Denote the Grassmann manifold of m-dimensional subspaces of $\mathbb{C}^n$ by $\mathscr{G}_{m,n}$ and the lattice of all subspaces of $\mathbb{C}^n$ by $\mathscr{G}_n$, i.e., $\mathscr{G}_n = \cup_{m=0}^n \mathscr{G}_{m,n}$. In this paper we will give a new family of unitarily invariant metrics on $\mathscr{G}_n$ based on the canonical angles. A function $\rho : \mathscr{G}_n \times \mathscr{G}_n \to \mathbb{R}$ is a metric if for any $\mathscr{X}, \mathscr{Y}, \mathscr{Z} \in \mathscr{G}_n$, we have

(i) $\rho(\mathscr{X}, \mathscr{Y}) \geqslant 0$; the equality holds if and only if $\mathscr{X} = \mathscr{Y}$,
(ii) $\rho(\mathscr{X}, \mathscr{Y}) = \rho(\mathscr{Y}, \mathscr{X})$,
(iii) $\rho(\mathscr{X}, \mathscr{Z}) \leqslant \rho(\mathscr{X}, \mathscr{Y}) + \rho(\mathscr{Y}, \mathscr{Z})$
and it is said to be unitarily invariant if
(iv) $\rho(U\mathscr{X}, U\mathscr{Y}) = \rho(\mathscr{X}, \mathscr{Y})$ for any unitary transformation U on $\mathbb{C}^n$.

In a recent paper [1], a family of unitarily invariant metrics is established in $\mathscr{G}_{m,n}$ by applying the symmetric gauge functions to the canonical angles. A *symmetric gauge function* $\Phi : \mathbb{R}^m \to \mathbb{R}$ is a norm function satisfying additional properties that it is symmetric, i.e.,

$$\Phi(P\xi) = \Phi(\xi)$$

for all $\xi \in \mathbb{R}^m$ and all permutation matrices P, and that it is absolute, i.e.,

$$\Phi(|\xi|) = \Phi(\xi)$$

for all $\xi \in \mathbb{R}^m$, where $|\xi|$ means the element-wise absolute value. A class of frequently used symmetric gauge function are given by

$$\Phi_{k,p}(\xi_1, \ldots, \xi_m) = \max_{1 \leqslant i_1 < \cdots < i_k \leqslant m} \sqrt[p]{|\xi_{i_1}|^p + \cdots + |\xi_{i_k}|^p}, \quad k = 1, \ldots, m, \quad p \in [1, \infty].$$

In particular, when $p = 1$, the above symmetric gauge functions are called the Ky Fan k-functions [2]. It is well-known that for a pair of vectors $\xi, \zeta \in \mathbb{R}^m$, if ζ *weakly majorizes* ξ, that is, $\Phi_{k,1}(\xi) \leqslant \Phi_{k,1}(\zeta)$ for all $k = 1, \ldots, m$, then we have $\Phi(\xi) \leqslant \Phi(\zeta)$ for all symmetric gauge functions Φ on $\mathbb{R}^m$.

The key results of paper [1] are listed below, and will be used in the proofs of the main results of this paper.

(I) Let $\Phi : \mathbb{R}^m \to \mathbb{R}$ be a symmetric gauge function. Define $\rho : \mathscr{G}_{m,n} \times \mathscr{G}_{m,n} \to \mathbb{R}$ by $\rho(\mathscr{X}, \mathscr{Y}) = \Phi(\theta(\mathscr{X}, \mathscr{Y}))$. Then ρ is a unitarily invariant metric and is called an angular matric.

(II) Let ρ be the angular metric corresponding to a symmetric gauge function Φ. Then for $\mathscr{X}, \mathscr{Y}, \widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}} \in \mathscr{G}_{m,n}$,

$$\Phi(\theta(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) - \theta(\mathscr{X}, \mathscr{Y})) \leqslant \rho(\mathscr{X}, \widetilde{\mathscr{X}}) + \rho(\mathscr{Y}, \widetilde{\mathscr{Y}}).$$

(III) For $\mathscr{X}, \mathscr{Y} \in \mathscr{G}_{m,n}$, define the nullity $\mathrm{nul}(\mathscr{X}, \mathscr{Y}) = \dim(\mathscr{X} \cap \mathscr{Y})$ and deficiency $\mathrm{def}(\mathscr{X}, \mathscr{Y}) = \mathrm{codim}(\mathscr{X} + \mathscr{Y})$. Again let ρ be the angular metric corresponding to a symmetric gauge function Φ. Then in order that $\mathrm{nul}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) < k$ for all $\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}} \in \mathscr{G}_{m,n}$ satisfying $\rho(\widetilde{\mathscr{X}}, \mathscr{X}) \leqslant \alpha$ and $\rho(\widetilde{\mathscr{Y}}, \mathscr{Y}) \leqslant \beta$, it is necessary and sufficient that

$$\alpha + \beta < \Phi[0, \ldots, 0, \theta_{m-k+1}(\mathscr{X}, \mathscr{Y}), \ldots, \theta_m(\mathscr{X}, \mathscr{Y})];$$

Authors personal copy

and in order that $\operatorname{def}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) < k$, it is necessary and sufficient that

$$\alpha + \beta < \Phi[0, \ldots, 0, \theta_{n-m-k+1}(\mathscr{X}_\perp, \mathscr{Y}_\perp), \ldots, \theta_{n-m}(\mathscr{X}_\perp, \mathscr{Y}_\perp)].$$

(IV) The angular metrics are intrinsic in the sense of [3], i.e., for each pair $\mathscr{X}, \mathscr{Y} \in \mathscr{G}_{m,n}$, there exists a curve (geodesics) $\mathscr{Z}_\lambda \subset \mathscr{G}_{m,n}$, $\lambda \in [0, 1]$, such that $\mathscr{Z}_0 = \mathscr{X}$, $\mathscr{Z}_1 = \mathscr{Y}$ and $\rho(\mathscr{X}, \mathscr{Y}) = \rho(\mathscr{X}, \mathscr{Z}_\lambda) + \rho(\mathscr{Z}_\lambda, \mathscr{Y})$ for all $\lambda \in [0, 1]$. This curve is actually given by a direct rotation from $\mathscr{X}$ and $\mathscr{Y}$ as defined in [4]. Let such a direct rotation be given by $\exp(A)$, where A is a skew-Hermitian matrix with eigenvalues in an imaginary interval $\left[-\frac{\pi}{2}i, \frac{\pi}{2}i\right]$. Then $\mathscr{Z}_\lambda = \exp(\lambda A)\mathscr{X}$.

We take this opportunity to conjecture that the angular metrics give all unitarily invariant *intrinsic* metrics on $\mathscr{G}_{m,n}$, but we will not further address this issue here. We argue that being intrinsic is the advantage of the angular metrics over other metrics. This property is essential for the inequality in result (II) to be sharp and it makes the result (III) possible.

In applications, it is often needed to carry out perturbation analysis involving subspaces of different dimensions. For example, the canonical angles are defined between subspaces of possibly different dimensions and the very perturbation analysis of the canonical angles involves subspaces of possibly different dimensions [5]. For another example, the related concepts of direct rotations and CS decompositions have the general versions involving subspaces of different dimensions [6,7]. Furthermore, the nullity and deficiency can also be defined for subspaces of possibly different dimensions, hence the robustness of nullity and deficiency between subspaces of different dimensions should be studied [8,9]. Finally, if one is to study all these notions in an infinite dimensional space setting, as in [4], then there is no reason to insist the subspaces involved to have equal dimension. All these motivates the need to extend the angular metrics to the space $\mathscr{G}_n$. When we define metrics on $\mathscr{G}_n$, the canonical angles cannot be directly used since they do not completely characterize the geometric relationship between two subspaces. For example, if $\mathscr{X} \subset \mathscr{Y}$ but $\mathscr{X} \neq \mathscr{Y}$, then $\theta(\mathscr{X}, \mathscr{Y}) = 0$ but it does not tell the excess of dimension of $\mathscr{Y}$ over that of $\mathscr{X}$ and the distance between $\mathscr{X}$ and $\mathscr{Y}$ cannot be zero.

A well-known family of metrics on $\mathscr{G}_n$ is the family of gap metrics ([10, p. 93] or [11, p. 202]) which is defined for $\mathscr{X}, \mathscr{Y} \in \mathscr{G}_n$ by

$$\delta(\mathscr{X}, \mathscr{Y}) = \|P_\mathscr{X} - P_\mathscr{Y}\|,$$

where $P_\mathscr{X}$ and $P_\mathscr{Y}$ are orthogonal projections onto $\mathscr{X}$ and $\mathscr{Y}$, respectively, and $\|\cdot\|$ is a unitarily invariant norm. More detailed examination reveals that if $\mathscr{X}$ is of dimension m and $\mathscr{Y}$ is of dimension l, the singular values of $P_\mathscr{X} - P_\mathscr{Y}$ are given by

$$1, \ldots, 1, \sin\theta_1, \sin\theta_1, \ldots, \sin\theta_{\min\{m,l,n-m,n-l\}}, \sin\theta_{\min\{m,l,n-m,n-l\}}, 0, \ldots, 0.$$

Here the number of extra 1's is $|l - m|$ and the number of extra zeros is determined so that the total number of singular values is n. We may view each 1 as $\sin\pi/2$ and each 0 as $\sin 0$. This suggests that in order to capture the missing information in the canonical angles between two subspaces, we may augment the canonical angles by certain number of $\pi/2$ and 0. The augmented canonical angles are exactly the quantities that we will use to define the angular metrics on $\mathscr{G}_n$.

When dealing with two subspaces of different dimensions, one may be interested in whether the subspace with smaller dimension is contained in or is closed to the one with larger dimension. Apparently, the metric between the two subspaces cannot be zero or close to zero. One may still wish to get the interested information from whatever metric he/she is using. Consider the gap metrics introduced above. If $\|\cdot\|$ is a unitarily invariant matrix norm corresponding to a strictly

Author's personal copy

monotone symmetric gauge function Φ, i.e., $\Phi(\xi) < \Phi(\zeta)$ if $|\xi| \leqslant |\zeta|$ but $|\xi| \neq |\zeta|$, where the absolute value and the inequality have the element-wise meaning, then for m-dimensional $\mathscr{X}$ and l-dimensional $\mathscr{Y}$ with $m < l$, we have $\mathscr{X} \subset \mathscr{Y}$ if and only if $\delta(\mathscr{X}, \mathscr{Y}) = \Phi(1, \ldots, 1, 0, \ldots, 0)$, where the number of 1's is equal to $l - m$, and $\mathscr{X}$ is close to $\mathscr{Y}$ if and only if the above inequality holds approximately. The metrics that we will define in this paper will also have the similar property.

In Section 2, we establish a family of metrics on $\mathscr{G}_n$ by applying the symmetric gauge functions to a set of properly augmented canonical angles, extending result (I) in [1]. We then study the perturbation of canonical angles between subspaces of unequal dimensions using the newly defined metrics, extending result (II) in [1]. In Section 3, we consider the robustness of nullity and deficiency between a pair of subspaces in $\mathscr{G}_n$, extending result (III) in [1]. Section 4 concludes the paper.

We will use $\mathscr{M}_{m,n}$ (respectively, $\mathscr{M}_n$) to denote the linear space of $m \times n$ (respectively, $n \times n$) matrices over $\mathbb{C}$. The group of unitary matrices is denoted by $\mathscr{U}_n$.

2. The extended angular metrics and perturbation of canonical angles

Let us take a closer look at the definition of canonical angles in (1) for the case when $m < l$. By the singular value decomposition, we know that there exist $E \in \mathscr{U}_m$ and $F \in \mathscr{U}_l$ such that

$$E^H X^H Y F = [\operatorname{diag}(\cos\theta_1(\mathscr{X}, \mathscr{Y}), \ldots, \cos\theta_m(\mathscr{X}, \mathscr{Y})) \quad 0_{m,l-m}].$$

This shows that $\mathscr{Y}$ has an m-dimensional subspace $\mathscr{Y}_0$ such that $\theta(\mathscr{X}, \mathscr{Y}_0) = \theta(\mathscr{X}, \mathscr{Y})$ and $\mathscr{X} \perp (\mathscr{Y} \ominus \mathscr{Y}_0)$. A similar interpretation can be made for the case when $l \leqslant m$. From the above observation, we may augment the angles between $\mathscr{X}$ and $\mathscr{Y}$ by adding $|l - m|$ angles of $\pi/2$. Since the length of $\theta(\mathscr{X}, \mathscr{Y})$ changes with the dimensions of $\mathscr{X}$ and $\mathscr{Y}$ as $\mathscr{X}$ and $\mathscr{Y}$ vary over $\mathscr{G}_n$, we also add some extra zeros to $\theta(\mathscr{X}, \mathscr{Y})$ so that the number of the augmented angles keeps invariant. We introduce the following n-tuple:

$$\bar{\theta}(\mathscr{X}, \mathscr{Y}) = \Big(\frac{\pi}{2}, \ldots, \frac{\pi}{2}, \theta_1(\mathscr{X}, \mathscr{Y}), \ldots, \theta_{\min\{m,l\}}(\mathscr{X}, \mathscr{Y}), 0, \ldots, 0\Big),$$

where the numbers of extra $\pi/2$ and 0 are $|l - m|$ and $n - \max\{m, l\}$, respectively. It is obvious that $\bar{\theta}(\mathscr{X}, \mathscr{Y})$ is unitarily invariant.

Using the augmented angles between subspaces, we are now able to extend the results (I) and (II) on $\mathscr{G}_{m,n}$ to $\mathscr{G}_n$.

Theorem 1. *Let $\Phi : \mathbb{R}^n \to \mathbb{R}$ be a symmetric gauge function. Define $\rho : \mathscr{G}_n \times \mathscr{G}_n \to \mathbb{R}$ by*

$$\rho(\mathscr{X}, \mathscr{Y}) = \Phi(\bar{\theta}(\mathscr{X}, \mathscr{Y})). \tag{2}$$

Then ρ is a unitarily invariant metric.

Proof. First $\rho(U\mathscr{X}, U\mathscr{Y}) = \rho(\mathscr{X}, \mathscr{Y})$ for all $U \in \mathscr{U}_n$, since $\bar{\theta}(\mathscr{X}, \mathscr{Y})$ is unitarily invariant from its definition. It is easy to see that $\rho(\mathscr{X}, \mathscr{Y}) = \rho(\mathscr{Y}, \mathscr{X})$, and $\rho(\mathscr{X}, \mathscr{Y}) \geqslant 0$. Moreover $\rho(\mathscr{X}, \mathscr{Y}) = 0$ implies that the entries of $\bar{\theta}(\mathscr{X}, \mathscr{Y})$ are all zeros, so the dimension of $\mathscr{X}$ equals that of $\mathscr{Y}$, and the canonical angles between them are all zeros, hence $\mathscr{X} = \mathscr{Y}$. Next we will prove the triangle inequality.

Let $X \in \mathscr{M}_{n,m}$, $Y \in \mathscr{M}_{n,l}$, and $Z \in \mathscr{M}_{n,t}$ have columns forming orthonormal bases for subspaces $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{Z}$, respectively. Without loss of generality, we assume that $m \leqslant l \leqslant t$. Define

Author's personal copy

$$\overline{X} = \begin{bmatrix} X & 0 \\ 0 & I_{n-m} \end{bmatrix}, \quad \overline{Y} = \begin{bmatrix} Y & 0 \\ 0_{l-m,l} & 0 \\ 0 & I_{n-l} \end{bmatrix}, \quad \overline{Z} = \begin{bmatrix} Z & 0 \\ 0_{t-m,t} & 0 \\ 0 & I_{n-t} \end{bmatrix}$$

and $\overline{\mathscr{X}}, \overline{\mathscr{Y}}, \overline{\mathscr{Z}} \in \mathscr{G}_{n,2n-m}$ to be subspaces spanned by the columns of $\overline{X}, \overline{Y}, \overline{Z}$, respectively. Then it is easy to see that the columns of $\overline{X}, \overline{Y}, \overline{Z}$ form orthonormal bases of $\overline{\mathscr{X}}, \overline{\mathscr{Y}}, \overline{\mathscr{Z}}$. Furthermore,

$$\overline{X}^H\overline{Y} = \begin{bmatrix} X^HY & 0 \\ 0 & 0 \\ 0 & I_{n-l} \end{bmatrix}, \quad \overline{X}^H\overline{Z} = \begin{bmatrix} X^HZ & 0 \\ 0 & 0 \\ 0 & I_{n-t} \end{bmatrix}, \quad \overline{Y}^H\overline{Z} = \begin{bmatrix} Y^HZ & 0 \\ 0 & 0 \\ 0 & I_{n-t} \end{bmatrix},$$

it follows from the definition of canonical angles that:

$$\theta(\overline{\mathscr{X}}, \overline{\mathscr{Y}}) = \bar{\theta}(\mathscr{X}, \mathscr{Y}), \quad \theta(\overline{\mathscr{X}}, \overline{\mathscr{Z}}) = \bar{\theta}(\mathscr{X}, \mathscr{Z}), \quad \theta(\overline{\mathscr{Y}}, \overline{\mathscr{Z}}) = \bar{\theta}(\mathscr{Y}, \mathscr{Z}). \tag{3}$$

Then by the result in (I), we have for any symmetric gauge function $\Phi : \mathbb{R}^n \to \mathbb{R}$,

$$\Phi(\bar{\theta}(\mathscr{X}, \mathscr{Z})) - \Phi(\bar{\theta}(\mathscr{Y}, \mathscr{Z})) \leqslant \Phi(\bar{\theta}(\mathscr{X}, \mathscr{Y})).$$

Hence

$$\rho(\mathscr{X}, \mathscr{Z}) \leqslant \rho(\mathscr{Y}, \mathscr{Z}) + \rho(\mathscr{X}, \mathscr{Y}).$$

This shows ρ is a unitarily invariant metric on $\mathscr{G}_n$. □

If the metrics defined above are restricted to Grassmannian $\mathscr{G}_{m,n}$, then they give exactly those metrics defined in [1].

By (3) and (II), we can easily get the following "Mirsky type result", which shows that the perturbations in $\bar{\theta}(\mathscr{X}, \mathscr{Y})$ are bounded by the perturbations in the subspaces involved.

Theorem 2. *Let $\Phi : \mathbb{R}^n \to \mathbb{R}$ be a symmetric gauge function, and ρ be the corresponding metric defined by (2). Then for any $\mathscr{X}, \mathscr{Y}, \mathscr{Z} \in \mathscr{G}_n$, we have*

$$\Phi(\bar{\theta}(\mathscr{X}, \mathscr{Z}) - \bar{\theta}(\mathscr{Y}, \mathscr{Z})) \leqslant \rho(\mathscr{X}, \mathscr{Y}).$$

In the case when both subspaces involved are perturbed, the following version of Theorem 2 might be more convenient in applications.

Corollary 3. *Let ρ be the metric corresponding to symmetric gauge function Φ on $\mathbb{R}^n$. Then for $\mathscr{X}, \mathscr{Y}, \widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}} \in \mathscr{G}_n$, we have*

$$\Phi(\bar{\theta}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) - \bar{\theta}(\mathscr{X}, \mathscr{Y})) \leqslant \rho(\widetilde{\mathscr{X}}, \mathscr{X}) + \rho(\widetilde{\mathscr{Y}}, \mathscr{Y}).$$

Proof. From the definition of symmetric gauge functions, it is easy to show that

$$\begin{aligned}\Phi(\bar{\theta}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) - \bar{\theta}(\mathscr{X}, \mathscr{Y})) &= \Phi(\bar{\theta}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) - \bar{\theta}(\mathscr{X}, \widetilde{\mathscr{Y}}) + \bar{\theta}(\mathscr{X}, \widetilde{\mathscr{Y}}) - \bar{\theta}(\mathscr{X}, \mathscr{Y})) \\ &\leqslant \Phi(\bar{\theta}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) - \bar{\theta}(\mathscr{X}, \widetilde{\mathscr{Y}})) + \Phi(\bar{\theta}(\mathscr{X}, \widetilde{\mathscr{Y}}) - \bar{\theta}(\mathscr{X}, \mathscr{Y})) \\ &\leqslant \rho(\widetilde{\mathscr{X}}, \mathscr{X}) + \rho(\widetilde{\mathscr{Y}}, \mathscr{Y}),\end{aligned}$$

where the last inequality holds from Theorem 2. □

In Theorem 2, if two of the three subspaces have same dimensions, we can use the canonical angles directly, instead of the augmented angles, to characterize the perturbations.

Author's personal copy

Corollary 4. *Let $\widehat{\Phi} : \mathbb{R}^m \to \mathbb{R}$ be a symmetric gauge function, and $\hat{\rho}$ be the corresponding angular metric on $\mathscr{G}_{m,n}$. Then for $\mathscr{X}, \mathscr{Y} \in \mathscr{G}_{m,n}$, $\mathscr{Z} \in \mathscr{G}_{l,n}$ with $m \leqslant l$, we have*

$$\widehat{\Phi}(\theta(\mathscr{X}, \mathscr{Z}) - \theta(\mathscr{Y}, \mathscr{Z})) \leqslant \hat{\rho}(\mathscr{X}, \mathscr{Y}).$$

3. Robustness of nullity and deficiency

In this section we will work on the issues related to the nullity and deficiency. For any two subspaces $\mathscr{X}, \mathscr{Y} \in \mathscr{G}_n$, following [12], the nullity and deficiency of $\mathscr{X}$ and $\mathscr{Y}$ are defined to be

$$\operatorname{nul}(\mathscr{X}, \mathscr{Y}) := \dim(\mathscr{X} \cap \mathscr{Y}) \quad \text{and} \quad \operatorname{def}(\mathscr{X}, \mathscr{Y}) := \operatorname{codim}(\mathscr{X} + \mathscr{Y}),$$

respectively. Clearly, in the case when $\dim \mathscr{X} + \dim \mathscr{Y} = n$, we have $\operatorname{def}(\mathscr{X}, \mathscr{Y}) = \operatorname{nul}(\mathscr{X}, \mathscr{Y})$. Also, $\mathscr{X} \oplus \mathscr{Y} = \mathbb{C}^n$ if and only if $\operatorname{nul}(\mathscr{X}, \mathscr{Y}) = 0$ and $\operatorname{def}(\mathscr{X}, \mathscr{Y}) = 0$.

The robustness of the nullity and deficiency is of great interest in mathematics [12], statistics [13], and control theory [8,9,14–16]. In particular, if we have subspaces $\mathscr{X} \in \mathscr{G}_{m,n}$ and $\mathscr{Y} \in \mathscr{G}_{l,n}$ with $\operatorname{nul}(\mathscr{X}, \mathscr{Y}) < k$ for some $1 \leqslant k \leqslant \min\{m, l\}$ and if we also know the perturbed versions $\widetilde{\mathscr{X}} \in \mathscr{G}_{m,n}$ and $\widetilde{\mathscr{Y}} \in \mathscr{G}_{l,n}$ satisfy $\rho(\mathscr{X}, \widetilde{\mathscr{X}}) \leqslant \alpha$ and $\rho(\mathscr{Y}, \widetilde{\mathscr{Y}}) \leqslant \beta$, we wish to obtain the tightest condition on α and β to ensure $\operatorname{nul}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) < k$. The same problem can be considered for the deficiency. The following theorem gives an answer.

Theorem 5. *Let $\Phi : \mathbb{R}^n \to \mathbb{R}$ be a symmetric gauge function, and ρ be the corresponding metric defined by (2). Let $\mathscr{X} \in \mathscr{G}_{m,n}$, $\mathscr{Y} \in \mathscr{G}_{l,n}$ with $m < l$. Then for $\alpha \geqslant 0$, $\beta \geqslant 0$, and $1 \leqslant k \leqslant m$,*

(1) *$\operatorname{nul}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) < k$ for all $\widetilde{\mathscr{X}} \in \mathscr{G}_{m,n}$ and $\widetilde{\mathscr{Y}} \in \mathscr{G}_{l,n}$ satisfying $\rho(\mathscr{X}, \widetilde{\mathscr{X}}) \leqslant \alpha$ and $\rho(\mathscr{Y}, \widetilde{\mathscr{Y}}) \leqslant \beta$ if and only if*
$$\alpha + \beta < \Phi(\theta_{m-k+1}(\mathscr{X}, \mathscr{Y}), \ldots, \theta_m(\mathscr{X}, \mathscr{Y}), 0, \ldots, 0);$$
(2) *$\operatorname{def}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) < k$ for all $\widetilde{\mathscr{X}} \in \mathscr{G}_{m,n}$ and $\widetilde{\mathscr{Y}} \in \mathscr{G}_{l,n}$ satisfying $\rho(\mathscr{X}, \widetilde{\mathscr{X}}) \leqslant \alpha$ and $\rho(\mathscr{Y}, \widetilde{\mathscr{Y}}) \leqslant \beta$ if and only if*
$$\alpha + \beta < \Phi(\theta_{n-l-k+1}(\mathscr{X}, \mathscr{Y}), \ldots, \theta_m(\mathscr{X}, \mathscr{Y}), 0, \ldots, 0).$$

Proof. We only need to prove Statement 1. Statement 2 follows from

$$\begin{aligned}\operatorname{def}(\mathscr{X}, \mathscr{Y}) &= n - \dim(\mathscr{X} + \mathscr{Y}) \\ &= n - \dim(\mathscr{X}) - \dim(\mathscr{Y}) + \dim(\mathscr{X} \cap \mathscr{Y}) \\ &= n - m - l + \operatorname{nul}(\mathscr{X}, \mathscr{Y}).\end{aligned}$$

Suppose $\operatorname{nul}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) \geqslant k$. Then $\theta_j(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) = 0$, $j = m - k + 1, \ldots, m$. By Corollary 3,

$$\begin{aligned}\delta &:= \Phi(\theta_{m-k+1}(\mathscr{X}, \mathscr{Y}), \ldots, \theta_m(\mathscr{X}, \mathscr{Y}), 0, \ldots, 0) \\ &\leqslant \Phi(\bar{\theta}(\mathscr{X}, \mathscr{Y}) - \bar{\theta}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}})) \leqslant \rho(\mathscr{X}, \widetilde{\mathscr{X}}) + \rho(\mathscr{Y}, \widetilde{\mathscr{Y}}) \leqslant \alpha + \beta.\end{aligned}$$

This shows that if

$$\alpha + \beta < \delta,$$

then $\operatorname{nul}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) < k$.

Now assume that $\alpha + \beta \geqslant \delta$. Then there exist $\alpha_1 \in [0, \alpha]$ and $\beta_1 \in [0, \beta]$ such that $\alpha_1 + \beta_1 = \delta$. Let $X_1 \in \mathscr{M}_{n,m}$ and $Y_1 \in \mathscr{M}_{n,l}$ be matrices whose columns form orthonormal bases of

Author's personal copy

$\mathscr{X}$ and $\mathscr{Y}$, respectively. By the singular value decomposition, we know that there exist $E_1 \in \mathscr{U}_m$ and $F_1 \in \mathscr{U}_l$ such that

$$E_1^H X_1^H Y_1 F_1 = [\operatorname{diag}(\cos\theta_1(\mathscr{X},\mathscr{Y}),\ldots,\cos\theta_m(\mathscr{X},\mathscr{Y})) \quad 0_{m,l-m}].$$

This shows that $\mathscr{Y}$ has an m-dimensional subspace $\mathscr{Y}_0$ such that $\theta(\mathscr{X},\mathscr{Y}_0)=\theta(\mathscr{X},\mathscr{Y})$ and $\mathscr{X}\perp(\mathscr{Y}\ominus\mathscr{Y}_0)$, hence there exist $X_2 \in \mathscr{M}_{n,n-l}$ and $Y_2 \in \mathscr{M}_{n,n-l}$ such that

$$X=\begin{bmatrix}X_1 & X_2 & Y_{12}\end{bmatrix},\quad Y=\begin{bmatrix}Y_{11} & Y_2 & Y_{12}\end{bmatrix}$$

are unitary matrices, where the columns of Y_{11} and Y_{12} form orthonormal bases of $\mathscr{Y}_0$ and $\mathscr{Y}\ominus\mathscr{Y}_0$, respectively. Let

$$Q=\begin{bmatrix}X_1^H\\ X_2^H\end{bmatrix}\begin{bmatrix}Y_{11} & Y_2\end{bmatrix}=\begin{bmatrix}X_1^HY_{11} & X_1^HY_2\\ X_2^HY_{11} & X_2^HY_2\end{bmatrix},$$

then it is easy to see that $Q \in \mathscr{U}_{n-l+m}$. In the following we assume $m+l<n$. For the case when $m+l \geqslant n$, the analysis is similar. By the CS decomposition, there are unitary matrices $U=\operatorname{diag}(U_{11},U_{22})$ and $V=\operatorname{diag}(V_{11},V_{22})$ with $U_{11}, V_{11} \in \mathscr{U}_m$ such that

$$U^HQV=\begin{bmatrix}(X_1U_{11})^H\\ (X_2U_{22})^H\end{bmatrix}\begin{bmatrix}Y_{11}V_{11} & Y_2V_{22}\end{bmatrix}=\begin{bmatrix}\Gamma & -\Sigma & 0\\ \Sigma & \Gamma & 0\\ 0 & 0 & I_{n-l-m}\end{bmatrix}, \tag{4}$$

where Γ and Σ are diagonal matrices with diagonal entries in $[0,1]$ satisfying $\Gamma^2+\Sigma^2=I_m$. Hence the diagonal entries of Γ are the cosines of the canonical angles between $\mathscr{X}$ and $\mathscr{Y}$; the diagonal entries of Σ are the sines of the canonical angles.

From (4), we may choose the columns of X_1, X_2, Y_{11}, Y_2 properly such that $U_{11}, U_{22}, V_{11}, V_{22}$ are identity matrices. We still use the notations X_1, X_2, Y_{11}, Y_2. Note that $Y_{12}^HY_{11}=0$ and X is unitary, we have

$$Y_{11}=\begin{bmatrix}X_1 & X_2\end{bmatrix}\begin{bmatrix}\cos(\operatorname{diag}\theta(\mathscr{X},\mathscr{Y}))\\ \sin(\operatorname{diag}\theta(\mathscr{X},\mathscr{Y}))\\ 0_{n-l-m,m}\end{bmatrix}. \tag{5}$$

Let

$$X_1=[x_1\cdots x_m],\quad X_2=[x_{m+1}\cdots x_{n-l+m}],$$
$$Y_{11}=[y_1\cdots y_m],\quad Y_{12}=[y_{n-l+m+1}\cdots y_n].$$

Next we will construct $\widetilde{\mathscr{X}} \in \mathscr{G}_{m,n}$ and $\widetilde{\mathscr{Y}} \in \mathscr{G}_{l,n}$. Let $\lambda=\alpha_1/\delta$, then $1-\lambda=\beta_1/\delta$. Also abbreviate $\theta_j(\mathscr{X},\mathscr{Y})$ by θ_j, $j=1,2,\ldots,m$. Define

$$\tilde{x}_j=x_j\cos(\lambda\theta_j)+x_{m+j}\sin(\lambda\theta_j),\quad j=m-k+1,\ldots,m$$

and

$$\widetilde{\mathscr{X}}=\operatorname{span}(x_1,\ldots,x_{m-k},\tilde{x}_{m-k+1},\ldots,\tilde{x}_m),$$
$$\widetilde{\mathscr{Y}}=\operatorname{span}(y_1,\ldots,y_{m-k},\tilde{x}_{m-k+1},\ldots,\tilde{x}_m,y_{n-l+m+1},\ldots,y_n).$$

Also define

$$\widetilde{X}_1=\begin{bmatrix}x_1 & \cdots & x_{m-k} & \tilde{x}_{m-k+1} & \cdots & \tilde{x}_m\end{bmatrix}$$
$$\widetilde{Y}_1=\begin{bmatrix}y_1 & \cdots & y_{m-k} & \tilde{x}_{m-k+1} & \cdots & \tilde{x}_m & y_{n-l+m+1} & \cdots & y_n\end{bmatrix}.$$

Author's personal copy

Then the columns of $\widetilde{X}_1$ and $\widetilde{Y}_1$ form orthonormal bases of $\widetilde{\mathscr{X}}$ and $\widetilde{\mathscr{Y}}$, respectively. Furthermore,

$$X_1^H \widetilde{X}_1 = \mathrm{diag}(1, \ldots, 1, \cos(\lambda\theta_{m-k+1}), \ldots, \cos(\lambda\theta_m)).$$

It follows that:

$$\rho(\mathscr{X}, \widetilde{\mathscr{X}}) = \Phi(\lambda\theta_{m-k+1}, \ldots, \lambda\theta_m, 0, \ldots, 0) = \alpha_1 \leqslant \alpha.$$

From (5), we see that $y_j = x_j \cos\theta_j + x_{m+j} \sin\theta_j,\ \ j = m-k+1, \ldots, m$. Hence

$$Y_1^H \widetilde{Y}_1 = \mathrm{diag}(1, \ldots, 1, \cos((1-\lambda)\theta_{m-k+1}), \ldots, \cos((1-\lambda)\theta_m)).$$

Similar to the above, we have

$$\rho(\mathscr{Y}, \widetilde{\mathscr{Y}}) = \Phi((1-\lambda)\theta_{m-k+1}, \ldots, (1-\lambda)\theta_m, 0, \ldots, 0) = \beta_1 \leqslant \beta.$$

Since

$$\widetilde{X}_1^H \widetilde{Y}_1 = [\mathrm{diag}(\cos\theta_1, \ldots, \cos\theta_{m-k}, 1, \ldots, 1)\ 0_{m,l-m}],$$

it follows that:

$$\theta(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) = (\theta_1, \ldots, \theta_{m-k}, 0, \ldots, 0)$$

which implies that $\mathrm{nul}(\widetilde{\mathscr{X}}, \widetilde{\mathscr{Y}}) \geqslant k$. This proves the necessity of the condition. □

4. Conclusions

In this work we have considered the metrics between subspaces with possibly different dimensions, and thus have extended the result in [1] to a more general setting. We have also studied the perturbation of canonical angles and the robustness of the nullity and the deficiency when the subspaces involved are perturbed using the newly defined metrics.

References

[1] L. Qiu, Y.X. Zhang, C.K. Li, Unitarily invariant metrics on the Grassmann space, SIAM J. Matrix Anal. Appl. 27 (2006) 507–531.
[2] K. Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, Proc. Natl. Acad. Sci. 37 (1951) 760–766.
[3] D. Burago, Y. Burago, S. Ivanov, A Course in Metric Geometry, American Mathematical Society, Providence, RI, 2001.
[4] C. Davis, W.M. Kahan, The rotation of eigenvectors by a perturbation. III, SIAM J. Numer. Anal. 7 (1970) 1–46.
[5] J.G. Sun, Perturbation of angles between linear subspaces, J. Comput. Math. 5 (1987) 58–61.
[6] C. Davis, W.M. Kahan, Some new bounds on perturbation of subspaces, Bull. Amer. Math. Soc. 75 (1969) 863–868.
[7] C.C. Paige, M. Wei, History and generality of the CS decomposition, Linear Algebra Appl. 208–209 (1994) 303–326.
[8] L. Qiu, E.J. Davison, Pointwise gap metrics on transfer matrices, IEEE Trans. Automat. Control 37 (1992) 741–758.
[9] H.K. Wimmer, Canonical angles of unitary spaces and perturbations of direct complements, Linear Algebra Appl. 287 (1999) 373–379.
[10] G.W. Stewart, J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[11] R. Bhatia, Matrix Analysis, Springer, New York, 1991.
[12] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, Berlin, New York, 1995.
[13] H. Hotelling, Relations between two sets of variates, Biometrika 28 (1936) 321–377.
[14] T.T. Georgiou, M.C. Smith, Optimal robustness in the gap metric, IEEE Trans. Automat. Control 35 (1990) 673–685.
[15] L. Qiu, E.J. Davison, Feedback stability under simultaneous gap metric uncertainties in plant and controller, Systems Control Lett. 18 (1992) 9–22.
[16] G. Vinnicombe, Frequency domain uncertainty and the graph topology, IEEE Trans. Automat. Control 38 (1993) 1371–1383.

Author's personal copy