SCIENCE CHINA
Mathematics

• **ARTICLES** •

December 2014 Vol. 57 No. 12: 2495–2508
doi: 10.1007/s11425-014-4824-0

Maximization of the sum of the trace ratio on the Stiefel manifold, I: Theory

ZHANG LeiHong[1] & LI RenCang[2,*]

[1]*Department of Applied Mathematics, Shanghai University of Finance and Economics, Shanghai* 200433*, China;*
[2]*Department of Mathematics, University of Texas at Arlington, Arlington, TX* 76019-0408*, USA*
Email: longzlh@163.com, rcli@uta.edu

Received July 30, 2013; accepted February 26, 2014; published online April 25, 2014

Abstract We are concerned with the maximization of

$$\frac{\mathrm{tr}(V^\top AV)}{\mathrm{tr}(V^\top BV)} + \mathrm{tr}(V^\top CV)$$

over the Stiefel manifold $\{V \in \mathbb{R}^{m\times\ell} \mid V^\top V = I_\ell\}$ $(\ell < m)$, where B is a given symmetric and positive definite matrix, A and C are symmetric matrices, and $\mathrm{tr}(\,\cdot\,)$ is the trace of a square matrix. This is a subspace version of the maximization problem studied in Zhang (2013), which arises from real-world applications in, for example, the downlink of a multi-user MIMO system and the sparse Fisher discriminant analysis in pattern recognition. We establish necessary conditions for both the local and global maximizers and connect the problem with a nonlinear extreme eigenvalue problem. The necessary condition for the global maximizers offers deep insights into the problem, on the one hand, and, on the other hand, naturally leads to a self-consistent-field (SCF) iteration to be presented and analyzed in detail in Part II of this paper.

Keywords trace ratio, Rayleigh quotient, Stiefel manifold, nonlinear eigenvalue problem, optimality condition, eigenspace

MSC(2010) 65F15, 65F30, 62H30, 15A18

Citation: Zhang L H, Li R C. Maximization of the sum of the trace ratio on the Stiefel manifold, I: Theory. Sci China Math, 2014, 57: 2495–2508, doi: 10.1007/s11425-014-4824-0

1 Introduction

We consider the maximization problem:

$$\max_{V^\top V = I_\ell} \left\{ \frac{\mathrm{tr}(V^\top AV)}{\mathrm{tr}(V^\top BV)} + \mathrm{tr}(V^\top CV) \right\}, \tag{1.1}$$

where $\mathrm{tr}(\,\cdot\,)$ stands for the trace of a square matrix, $A, B, C \in \mathbb{R}^{m\times m}$ are real symmetric with B positive definite, and integer $\ell < m$.

The problem (1.1) for the case $\ell = 1$ was investigated in [14]. It can arise from real-world applications in, e.g., the downlink of a multi-user MIMO system [9] and the sparse Fisher discriminant analysis in pattern recognition [3, 5, 7, 14]. In [14], the optimality conditions for (1.1) with $\ell = 1$, including necessary conditions for local and global maximizers, are established. It is interesting to note that these conditions

*Corresponding author

connect the optimization problem (1.1) with $\ell = 1$ to a nonlinear eigenvalue problem which offers fresh insights into (1.1) and injects new ideas for designing efficient algorithms.

The present paper attempts to deal with the subspace version (1.1) and generalize these interesting properties for $\ell = 1$ to $\ell > 1$. But we point out that although the problem (1.1) with $\ell > 1$ and $\ell = 1$ share certain similar properties, the subspace version case is intrinsically harder, entailing new techniques for investigation.

This is the first part of ours in a sequel on the subject. Here we focus on treating the theoretical aspect of the maximization problem (1.1), and its numerical aspect will be the subject of study in [16]. We begin by the standard first order optimality condition to conclude that any critical point (i.e., the KKT point) V of (1.1) is a solution to a nonlinear eigenvalue problem

$$E(V)V = V[V^\top E(V)V], \tag{1.2}$$

where $E(V) \in \mathbb{R}^{m\times m}$ to be given in Section 2 is symmetric and depends on V. Our major contributions here are the following:

(1) A necessary condition for the critical point V to be a local maximizer is that the largest eigenvalue of $V^\top E(V)V$ is no smaller than the (2ℓ)-th largest eigenvalues of $E(V)$. (Note the eigenvalues of $V^\top E(V)V$ are part of those of $E(V)$ because of (1.2).) (2) A necessary condition for the critical point V to be a global maximizer is that V is an orthonormal eigenbasis [1)] of $E(V)$ corresponding to its ℓ largest eigenvalues. This necessary condition together with (1.2) lend themselves to a self-consistent-field (SCF) iteration for computing a global maximizer. (3) We propose a scheme to march out a KKT point V, local maximizer included, that does not satisfy the necessary condition of a global maximizer in the previous item.

The rest of this paper is organized as follows. Section 2 develops the first and second order optimality conditions for local maximizers using the traditional optimality conditions which relate (1.1) to a special nonlinear eigenvalue problem (1.2). In Sections 3 and 4, further characterizations of a local maximizer (Section 3) and global maximizer (Section 4) in terms of the eigenvalues of the associated nonlinear eigenvalue problem are obtained. A scheme to break out of a local maximizer/KKT point is proposed in Section 5. Finally, we present our conclusions in Section 6.

Notation. $\mathbb{R}^{n\times m}$ is the set of all $n \times m$ real matrices, $\mathbb{R}^n = \mathbb{R}^{n\times 1}$, and $\mathbb{R} = \mathbb{R}^1$. I_n (or simply I if its dimension is clear from the context) is the $n \times n$ identity matrix. All vectors are column vectors and are in bold. For $Z \in \mathbb{R}^{m\times n}$, $Z^\top$ denotes its transpose, $\mathcal{R}(Z)$ is its column space, spanned by its column vectors. For $Z \in \mathbb{R}^{m\times m}$, $\mathrm{eig}(Z) = \{\lambda_i(Z),\, 1 \leqslant i \leqslant m\}$ is the set of the eigenvalues of Z, and if also all $\lambda_i(Z)$ are real, they will be arranged in descending order:

$$\lambda_1(Z) \geqslant \lambda_2(Z) \geqslant \cdots \geqslant \lambda_m(Z).$$

The symmetric part of Z is $\mathrm{sym}(Z) := \frac{1}{2}(Z + Z^\top)$.

2 First and second optimality conditions

To simplify our presentation, we introduce the following notation,

$$\phi_A(V) := \mathrm{tr}(V^\top AV), \quad \phi_B(V) := \mathrm{tr}(V^\top BV), \quad \phi_C(V) := \mathrm{tr}(V^\top CV), \tag{2.1}$$

for any $V \in \mathbb{R}^{m\times\ell}$. The objective function in (1.1) now can be written as

$$f(V) := \frac{\phi_A(V)}{\phi_B(V)} + \phi_C(V), \tag{2.2}$$

subject to the constraint $V^\top V = I_\ell$. Let

$$\mathbb{O}^{m\times\ell} := \{V \in \mathbb{R}^{m\times\ell} \,:\, V^\top V = I_\ell\}.$$

[1)] We follow [11, p. 242] in using the term *eigenbasis* to refer to a matrix X whose columns are linearly independent and span an invariant subspace of another matrix H. X is an *orthonormal eigenbasis* if its columns are orthonormal vectors.

We remark that maximizing $f(V)$ over $\mathbb{O}^{m\times\ell}$ is much difficult than maximizing each term, i.e., $\frac{\phi_A(V)}{\phi_B(V)}$ and $\phi_C(V)$ individually on $\mathbb{O}^{m\times\ell}$. A classical result for

$$\max_{V\in\mathbb{O}^{m\times\ell}} \operatorname{tr}(V^\top CV) \tag{2.3}$$

reveals that any solution V of (2.3) is an orthonormal eigenbasis associated with the ℓ largest eigenvalues of C and there is no *local but non-global* maximizer (i.e., any local maximizer is also a global maximizer). A recent result in [10] (see also [17, Theorem 1.1]) for the trace ratio problem

$$\max_{V\in\mathbb{O}^{m\times\ell}} \frac{\operatorname{tr}(V^\top AV)}{\operatorname{tr}(V^\top BV)} \tag{2.4}$$

shows that no local but non-global maximizer exists, either. This nice property ensures that any monotonic and convergent iteration is capable of finding a global maximizer for (2.3) or (2.4) *numerically*. However, when $f(V)$ is to be maximized over $\mathbb{O}^{m\times\ell}$, the story is very different because local but non-global maximizers can happen as a numerical example for the case $\ell = 1$ in [14, Example 3.1] shows.

Often necessary and sufficient conditions for global maximizers of an optimization problem are generally difficult to establish. This is the case for maximizing $f(V)$ here, too. In this section, we will first resort to the traditional approach to derive first order and second order optimality conditions. They usually follow from the classical Lagrange multiplier theory (e.g., [8]), but considering the nice geometry properties of the underlying Riemannian manifold $\mathbb{O}^{m\times\ell}$, we prefer to treating the function $f(V)$ as the restriction

$$f_{|\mathbb{O}^{m\times\ell}}(V) : \mathbb{O}^{m\times\ell} \to \mathbb{R}$$

that is to be maximized as an *unconstrained* optimization problem. For more discussions on the optimality conditions of the nonlinear programming on manifolds, the reader is referred to [1,13] and the references therein.

We can view $\mathbb{O}^{m\times\ell}$ as an embedded submanifold (the *Stiefel manifold*, see e.g., [1,2,4,6]) of the Euclidean space $\mathbb{R}^{m\times\ell}$. The tangent space $\mathcal{T}_V\mathbb{O}^{m\times\ell}$ at any $V \in \mathbb{O}^{m\times\ell}$ can be expressed as (see [1, p. 42] and [2, Theorem 1])

$$\begin{aligned}\mathcal{T}_V\mathbb{O}^{m\times\ell} &:= \{X \in \mathbb{R}^{m\times\ell} : X^\top V + V^\top X = 0\} && (2.5a)\\ &= \{X = VK + (I_m - VV^\top)J : K = -K^\top \in \mathbb{R}^{\ell\times\ell}, J \in \mathbb{R}^{m\times\ell}\}. && (2.5b)\end{aligned}$$

On $\mathcal{T}_V\mathbb{O}^{m\times\ell}$, we introduce the standard inner product (or the Frobenius inner product)

$$\langle X, Y\rangle = \operatorname{tr}(X^\top Y), \quad \forall\, X, Y \in \mathcal{T}_V\mathbb{O}^{m\times\ell}.$$

It is known that the orthogonal projection of $Z \in \mathbb{R}^{m\times\ell}$ onto the tangent space $\mathcal{T}_V\mathbb{O}^{m\times\ell}$ at V is given by (see [1, (3.35)] and [2, Corollary 1])

$$\begin{aligned}\Pi_{\mathcal{T}}(Z) &:= V\left(\frac{V^\top Z - Z^\top V}{2}\right) + (I_m - VV^\top)Z && (2.6a)\\ &= Z - V\frac{V^\top Z + Z^\top V}{2} = Z - V\,\operatorname{sym}(V^\top Z) \in \mathcal{T}_V\mathbb{O}^{m\times\ell}. && (2.6b)\end{aligned}$$

2.1 First order optimality condition

Using the projection defined by (2.6), we can find the gradient of $f_{|\mathbb{O}^{m\times\ell}}(V)$. In fact,

$$\frac{\partial f(V)}{\partial V} = 2\underbrace{\left[A\frac{1}{\phi_B(V)} - B\frac{\phi_A(V)}{[\phi_B(V)]^2} + C\right]}_{=:E(V)} V = 2\,E(V)V. \tag{2.7}$$

Thus the gradient $\operatorname{grad} f_{|\mathbb{O}^{m\times\ell}}(V)$ is given by (see e.g., [1, (3.37)])

$$\operatorname{grad} f_{|\mathbb{O}^{m\times\ell}}(V) = \Pi_{\mathcal{T}}\left(\frac{\partial f(V)}{\partial V}\right) = 2\{E(V)V - V\underbrace{[V^\top E(V)V]}_{=:M_V}\}. \tag{2.8}$$

The first order optimality condition $\operatorname{grad} f_{|\mathbb{O}^{m\times\ell}}(V) = \mathbf{0}$ leads to

Theorem 2.1. *If $V \in \mathbb{O}^{m\times\ell}$ is a local maximizer of* (1.1)*, then*

$$E(V)V = VM_V, \tag{2.9}$$

where $E(V)$ is defined in (2.7)*, and M_V is defined in* (2.8)*. Therefore* $\operatorname{eig}(M_V) \subset \operatorname{eig}(E(V))$*, and V is an orthonormal eigenbasis of $E(V)$ associated with its eigenvalues given by* $\operatorname{eig}(M_V)$*.*

Remark 2.2. Under the condition of Theorem 2.1, $\operatorname{tr}(M_V) = \phi_C(V)$. Since $f(VZ) \equiv f(V)$ for any orthogonal matrix $Z \in \mathbb{R}^{\ell\times\ell}$, if V is a local maximizer, so is VZ. That is any local maximizer V gives rise to a set of local maximizers $\{VZ : Z^\top Z = I_\ell\}$ and they all produce the same objective value. Within the set, there is one V_0 that makes M_{V_0} diagonal:

$$M_{V_0} = \operatorname{diag}(\lambda_{\pi_1}(E(V_0)), \dots, \lambda_{\pi_\ell}(E(V_0))), \quad V_0 = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell], \tag{2.10}$$

where $\boldsymbol{v}_i \in \mathbb{R}^m$ is the unit eigenvector corresponding to $\lambda_{\pi_i}(E(V_0))$, and

$$\lambda_{\pi_1}(E(V_0)) \geqslant \cdots \geqslant \lambda_{\pi_\ell}(E(V_0)).$$

This observation is also reflected by the equation (2.9) which is equivalent to

$$E(VZ)VZ = VZM_{VZ},$$

for any given orthogonal matrix $Z \in \mathbb{R}^{\ell\times\ell}$.

2.2 Second order optimality condition

Next, we will establish the second order optimality condition for (1.1). Set

$$g(V) := \operatorname{grad} f_{|\mathbb{O}^{m\times\ell}}(V).$$

Assume $V \in \mathbb{O}^{m\times\ell}$ is a critical point, i.e., V satisfies (2.9), throughout this subsection. By using the standard second order optimality conditions [8, Theorems 12.5 and 12.6], which can be reformulated simply as [12, Lemma 2] for the Stiefel manifold constraint, we have the following theorem.

Theorem 2.3. *If V is a local maximizer of* (1.1)*, then*

$$\operatorname{tr}(X^\top E(V)X) - \operatorname{tr}(XM_VX^\top) + \operatorname{tr}(X^\top G(V,X)V) \leqslant 0 \quad \text{for any} \quad X \in \mathcal{T}_V\mathbb{O}^{m\times\ell}, \tag{2.11}$$

where

$$G(V,X) := 4\frac{\operatorname{tr}(V^\top AV)\operatorname{tr}(X^\top BV)}{[\operatorname{tr}(V^\top BV)]^3}B - 2\frac{\operatorname{tr}(X^\top BV)A + \operatorname{tr}(X^\top AV)B}{[\operatorname{tr}(V^\top BV)]^2} \tag{2.12}$$

and $M_V = V^\top E(V)V$ in (2.8)*. On the other hand, if $V \in \mathbb{O}^{m\times\ell}$ satisfies* (2.9) *and if* (2.11) *is a strict inequality for $X \neq 0$, then V is a strict local maximizer.*

Proof. We first calculate

$$\mathscr{D}(\mathscr{D}f(V))[X] = 2\mathscr{D}(E(V)V)[X] = 2E(V)X + 2\mathscr{D}E(V)[X]V = 2E(V)X + 2G(V,X)V,$$

where

$$\begin{aligned}\mathscr{D}E(V)[X] &= -\frac{2}{[\operatorname{tr}(V^\top BV)]^2}[\operatorname{tr}(X^\top BV)A + \operatorname{tr}(X^\top AV)B] + \frac{4\operatorname{tr}(V^\top AV)\operatorname{tr}(X^\top BV)}{[\operatorname{tr}(V^\top BV)]^3}B \\ &= 4\frac{\operatorname{tr}(V^\top AV)\operatorname{tr}(X^\top BV)}{[\operatorname{tr}(V^\top BV)]^3}B - 2\frac{\operatorname{tr}(X^\top BV)A + \operatorname{tr}(X^\top AV)B}{[\operatorname{tr}(V^\top BV)]^2} \\ &=: G(V,X).\end{aligned}$$

Moreover, if V is a critical point, $V^\top \frac{\partial f(V)}{\partial V} = 2M_V$ and therefore, by [12, Lemma 2], the necessary second order optimality condition is

$$2\operatorname{tr}(X^\top E(X)X) + 2\operatorname{tr}(X^\top G(X,V)V) - 2\operatorname{tr}(M_V X^\top X) \leqslant 0, \quad \forall X \in \mathcal{T}_V \mathbb{O}^{m\times\ell},$$

which is (2.11), and consequently, the sufficient condition follows directly. A more self-contained proof based on Riemannian Hessian operator is given in [15]. □

The second order optimality condition in Theorem 2.3 is presented in terms of the tangent vector $X \in \mathcal{T}_V \mathbb{O}^{m\times\ell}$. In the next theorem, we describe the optimality condition in terms of any $J \in \mathbb{R}^{m\times\ell}$.

Theorem 2.4. *If V is a local maximizer of (1.1), then for all $J \in \mathbb{R}^{m\times\ell}$,*

$$\begin{aligned}&\operatorname{tr}(J^\top E(V)J) + \operatorname{tr}(V^\top J M_V J^\top V) - \operatorname{tr}(J^\top V M_V V^\top J) - \operatorname{tr}(J M_V J^\top)\\ &\quad + 4\frac{\operatorname{tr}(J^\top [I_m - VV^\top]BV)\operatorname{tr}(J^\top [I_m - VV^\top]CV)}{\phi_B(V)} \leqslant 0. \end{aligned} \tag{2.13}$$

On the other hand, if $V \in \mathbb{O}^{m\times\ell}$ satisfies (2.9) and if (2.13) is strict for any nonzero $J \in \mathbb{R}^{m\times\ell}$, then V is a strict local maximizer.

Proof. By (2.5b), we know that

$$\mathcal{T}_V \mathbb{O}^{m\times\ell} = \{X = VK + (I_m - VV^\top)J \,:\, K = -K^\top \in \mathbb{R}^{\ell\times\ell}, J \in \mathbb{R}^{m\times\ell}\},$$

i.e., $X = VK + (I_m - VV^\top)J \in \mathcal{T}_V \mathbb{O}^{m\times\ell}$ for any skew-symmetric K and arbitrary $J \in \mathbb{R}^{m\times\ell}$. For such an X, we have

$$E(V)X = E(V)VK + E(V)J - E(V)VV^\top J = VM_V K + E(V)J - VM_V V^\top J, \tag{2.14}$$

where we have used (2.9), and

$$\begin{aligned}\operatorname{tr}(X^\top E(V)X) &= \operatorname{tr}(K^\top M_V K) + \operatorname{tr}(K^\top V^\top E(V)J)\\ &\quad - \operatorname{tr}(K^\top M_V V^\top J) + \operatorname{tr}(J^\top [I_m - VV^\top]E(V)J)\\ &= \operatorname{tr}(K^\top M_V K) + \operatorname{tr}(J^\top [I_m - VV^\top]E(V)J)\\ &= \operatorname{tr}(K^\top M_V K) + \operatorname{tr}(J^\top E(V)J) - \operatorname{tr}(J^\top V M_V V^\top J).\end{aligned}$$

It can be verified that $X^\top X = K^\top K + J^\top (I_m - VV^\top)J$. Thus

$$\begin{aligned}\operatorname{tr}(X^\top X M_V) &= \operatorname{tr}(K^\top K M_V) + \operatorname{tr}(J^\top J M_V) - \operatorname{tr}(J^\top VV^\top J M_V)\\ &= \operatorname{tr}(K M_V K^\top) + \operatorname{tr}(J^\top J M_V) - \operatorname{tr}(V^\top J M_V J^\top V)\\ &= \operatorname{tr}(K^\top M_V K) + \operatorname{tr}(J^\top J M_V) - \operatorname{tr}(V^\top J M_V J^\top V),\end{aligned}$$

where we have used $K = -K^\top$ for the last equality. Hence,

$$\begin{aligned}&\operatorname{tr}(X^\top E(V)X) - \operatorname{tr}(X M_V X^\top)\\ &\quad = \operatorname{tr}(J^\top E(V)J) + \operatorname{tr}(V^\top J M_V J^\top V) - \operatorname{tr}(J^\top V M_V V^\top J) - \operatorname{tr}(J M_V J^\top).\end{aligned} \tag{2.15}$$

Lastly, we compute $\operatorname{tr}(X^\top G(V,X)V)$ as follows,

$$\begin{aligned}\operatorname{tr}(X^\top G(V,X)V) &= \frac{4\phi_A(V)[\operatorname{tr}(X^\top BV)]^2}{[\phi_B(V)]^3} - \frac{4\operatorname{tr}(X^\top BV)\operatorname{tr}(X^\top AV)}{\phi_B^2(V)}\\ &= \frac{4\operatorname{tr}(X^\top BV)}{[\phi_B(V)]^3}[\phi_A(V)\operatorname{tr}(X^\top BV) - \phi_B(V)\operatorname{tr}(X^\top AV)]\\ &= \frac{4\operatorname{tr}(X^\top BV)}{[\phi_B(V)]^3}\langle X, [\phi_A(V)B - \phi_B(V)A]V\rangle\end{aligned}$$

$$= \frac{4\,\mathrm{tr}(X^\top BV)}{[\phi_B(V)]^3}\langle X, [\phi_B(V)]^2[C - E(V)]V\rangle$$
$$= \frac{4\,\mathrm{tr}(X^\top BV)}{\phi_B(V)}\langle X, [C - E(V)]V\rangle. \tag{2.16}$$

Note from (2.14) that

$$\begin{aligned}\langle X, E(V)V\rangle &= \langle V, VM_VK + E(V)J - VM_VV^\top J\rangle \\ &= \mathrm{tr}(M_VK) + \mathrm{tr}(J^\top E(V)V) - \mathrm{tr}(M_VV^\top J) \\ &= \mathrm{tr}(J^\top VM_V) - \mathrm{tr}(M_VV^\top J) = 0,\end{aligned} \tag{2.17}$$

and

$$\langle X, CV\rangle = \langle VK + (I_m - VV^\top)J, CV\rangle = \mathrm{tr}(J^\top[I_m - VV^\top]CV),$$
$$\langle X, BV\rangle = \langle VK + (I_m - VV^\top)J, BV\rangle = \mathrm{tr}(J^\top[I_m - VV^\top]BV),$$

which, together with (2.11) and (2.15)–(2.17), lead to (2.13). The second part of this theorem follows directly from Theorem 2.3. □

3 A necessary condition for local maximizers

The purpose of this section is to establish a necessary condition for a local maximizer V of (1.1) in terms of the eigenvalues of $E(V)$. The same will be done for a global maximizer in Section 4.

Suppose V is a local maximizer of (1.1). According to Theorem 2.1 and Remark 2.2, $\mathrm{eig}(M_V) \subset \mathrm{eig}(E(V))$, i.e.,

$$\mathrm{eig}(M_V) = \{\lambda_{\pi_i}(E(V)),\ i = 1, 2, \ldots, \ell\}, \tag{3.1}$$

where $1 \leqslant \pi_1 < \pi_2 \cdots < \pi_\ell \leqslant m$.

Theorem 3.1. *Let $V \in \mathbb{O}^{m\times\ell}$ be a local maximizer of* (1.1), *and denote* $\mathrm{eig}(M_V)$ *by* (3.1). *Then*

$$\lambda_{\pi_1}(E(V)) \geqslant \lambda_{2\ell}(E(V)). \tag{3.2}$$

Proof. Assume, to the contrary, that (3.2) were false, i.e.,

$$\lambda_{\pi_1}(E(V)) < \lambda_{2\ell}(E(V)). \tag{3.3}$$

Let $J_1 \in \mathbb{R}^{m\times 2\ell}$ be an orthonormal eigenbasis of $E(V)$ corresponding to its 2ℓ largest eigenvalues $\lambda_i(E(V))$ for $1 \leqslant i \leqslant 2\ell$. Then $J_1^\top V = 0$ because V is the orthonormal eigenbasis of $E(V)$ corresponding to its eigenvalues $\lambda_{\pi_i}(E(V))$ for $1 \leqslant i \leqslant \ell$, all of which were assumed less than $\lambda_{2\ell}(E(V))$ by (3.3). For any $Q \in \mathbb{R}^{2\ell\times\ell}$ with orthonormal columns, let $J = J_1Q \in \mathbb{R}^{m\times\ell}$. Then $J^\top J = I_\ell$ and $J^\top V = Q^\top J_1^\top V = 0$, and thus

$$\mathrm{tr}(J^\top E(V)J) > \mathrm{tr}(V^\top E(V)V) = \mathrm{tr}(M_V) = \mathrm{tr}(J^\top JM_V) = \mathrm{tr}(JM_VJ^\top). \tag{3.4}$$

Therefore, it follows from (2.13) that

$$\mathrm{tr}(J^\top E(V)J) - \mathrm{tr}(M_V) + 4\frac{\mathrm{tr}(J^\top BV)\,\mathrm{tr}(J^\top CV)}{\phi_B(V)} \leqslant 0. \tag{3.5}$$

On the other hand, we note that

$$J^\top BV = Q^\top(J_1^\top BV) = Q^\top(\widehat{U}\varSigma_1\widehat{W}^\top),$$

where $\widehat{U}\varSigma_1\widehat{W}^\top$ is the SVD of $J_1^\top BV$ and $\widehat{U} \in \mathbb{R}^{2\ell\times\ell}$, $\varSigma_1 \in \mathbb{R}^{\ell\times\ell}$ and $\widehat{W} \in \mathbb{R}^{\ell\times\ell}$. Since Q is arbitrary so far, we now let Q be the one such that $[\widehat{U}, Q] \in \mathbb{R}^{2\ell\times 2\ell}$ is orthogonal. We have

$$J^\top BV = Q^\top(J_1^\top BV) = Q^\top\widehat{U}\varSigma_1\widehat{W}^\top = 0$$

which leads to

$$\operatorname{tr}(J^\top E(V)J) - \operatorname{tr}(M_V) \leqslant 0,$$

contradicting (3.4). Therefore (3.3) cannot be true. That gives (3.2). □

Theorem 3.1 basically says that any local maximizer V will never be an orthonormal eigenbasis of $E(V)$ associated with the ℓ eigenvalues all less than $\lambda_{2\ell}(E(V))$. For the special case with $\ell = 1$, Theorem 3.1 implies that any local maximizer $\boldsymbol{v} \in \mathbb{R}^m$ can only be the eigenvector associated with the largest or the second largest eigenvalue of $E(\boldsymbol{v})$. This is exactly what was proved in [14, Theorem 3.3], even though the matrix $E(\boldsymbol{v})$ in [14] takes a different form. Theorem 3.1 can be considered as a generalization of [14, Theorem 3.3] for the necessary condition of a local maximizer.

4 A necessary condition for global maximizers

In this section and the next, C is assumed to be positive definite. This assumption is not essential because one can always shift C to $C + \xi I_m$ for some $\xi \in \mathbb{R}$ to make C positive definite. The shift does not change the maximizer V.

Theorem 4.1. *Suppose C is positive definite. If V is a global maximizer of* (1.1)*, then it must be an orthonormal eigenbasis* [2] *of $E(V)$ corresponding to its ℓ largest eigenvalues $\lambda_i(E(V))$ for $1 \leqslant i \leqslant \ell$.*

This result generalizes the necessary condition for the vector version [14, Theorem 3.5] of the problem (1.1), i.e., the case $\ell = 1$. Its proof, similar to [14], hinges on the relationship between

$$\varDelta f(J,V) := f(J) - f(V) \quad \text{and} \quad \varDelta E(J,V) := \operatorname{tr}(J^\top E(V)J) - \operatorname{tr}(V^\top E(V)V). \tag{4.1}$$

This is established in the following lemma.

Lemma 4.2. *For any $J, V \in \mathbb{O}^{m\times\ell}$, we have*

$$\begin{aligned}\varDelta f(J,V) &= f(J) - f(V)\\ &= \frac{\phi_B(V)\varDelta E(J,V) + [\phi_B(J) - \phi_B(V)][\phi_C(J) - \phi_C(V)]}{\phi_B(J)}.\end{aligned} \tag{4.2}$$

Proof. Note that $\operatorname{tr}(V^\top E(V)V) = \phi_C(V)$ and

$$\begin{aligned}\varDelta E(J,V) &= \frac{\phi_A(J)}{\phi_B(V)} - \frac{\phi_A(V)\phi_B(J)}{[\phi_B(V)]^2} + \phi_C(J) - \phi_C(V)\\ &= \frac{\phi_B(J)}{\phi_B(V)}\left(\frac{\phi_A(J)}{\phi_B(J)} - \frac{\phi_A(V)}{\phi_B(V)}\right) + \phi_C(J) - \phi_C(V)\\ &= \frac{\phi_B(J)}{\phi_B(V)}[\varDelta f(J,V) - \phi_C(J) + \phi_C(V)] + \phi_C(J) - \phi_C(V)\\ &= \frac{\phi_B(J)}{\phi_B(V)}\varDelta f(J,V) + \frac{\phi_B(V) - \phi_B(J)}{\phi_B(V)}[\phi_C(J) - \phi_C(V)]\end{aligned}$$

solving which for $\varDelta f(J,V)$ leads to (4.2). □

Rewrite (4.2) as

$$\varDelta f(J,V) = f(J) - f(V) = \frac{\phi_B(V)\varDelta E(J,V) + \delta(J,V)}{\phi_B(J)}, \tag{4.3}$$

where

$$\delta(J,V) := [\phi_B(J) - \phi_B(V)][\phi_C(J) - \phi_C(V)]. \tag{4.4}$$

We are now ready to prove Theorem 4.1.

[2] $\mathcal{R}(V)$ is not unique unless $\lambda_\ell(E(V)) > \lambda_{\ell+1}(E(V))$.

Proof of Theorem 4.1. We prove it by contradiction. Suppose $V = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell]$ is a global maximizer but is not an orthonormal eigenbasis associated with the ℓ largest eigenvalues of $E(V)$. We shall construct $U \in \mathbb{O}^{m\times\ell}$ such that $f(U) > f(V)$, contradicting the assumption that V is a global maximizer. So the conclusion of the theorem holds.

As we commented in Remark 2.2, we may assume, without loss of generality, that

$$E(V)\boldsymbol{v}_i = \lambda_{\pi_i}(E(V))\boldsymbol{v}_i, \quad i = 1, 2, \ldots, \ell,$$

where $(\lambda_{\pi_i}(E(V)), \boldsymbol{v}_i)$ is the π_i-th largest eigenpair of $E(V)$ and

$$\lambda_{\pi_1}(E(V)) \geqslant \lambda_{\pi_2}(E(V)) \geqslant \cdots \geqslant \lambda_{\pi_\ell}(E(V)).$$

According to our assumption, we know that there exists at least one i such that

$$\lambda_{\pi_i}(E(V)) < \lambda_i(E(V)).$$

Let p be the first such an i, i.e.,

$$p := \min_{1\leqslant i\leqslant \ell}\{i : \lambda_{\pi_i}(E(V)) < \lambda_i(E(V))\}. \tag{4.5}$$

There is a unit eigenvector $\boldsymbol{u}$ of $E(V)$ such that [3)]

$$E(V)\boldsymbol{u} = \lambda_p(E(V))\boldsymbol{u}, \quad \text{and} \quad \boldsymbol{u}^\top \boldsymbol{v}_i = 0 \quad \text{for } i = 1, 2, \ldots, \ell. \tag{4.6}$$

Now, for any $\alpha, \beta \in \mathbb{R}$, define

$$\begin{aligned} U_{\alpha,\beta} &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{p-1}, \alpha\boldsymbol{u} + \beta\boldsymbol{v}_p, \boldsymbol{v}_{p+1}, \ldots, \boldsymbol{v}_\ell] \\ &= V - \underbrace{[(1-\beta)\boldsymbol{v}_p - \alpha\boldsymbol{u}]}_{=:\boldsymbol{y}}\boldsymbol{e}_p^\top \\ &= V - \boldsymbol{y}\boldsymbol{e}_p^\top \in \mathbb{R}^{m\times\ell}, \end{aligned} \tag{4.7}$$

where $\boldsymbol{e}_p$ is the p-th column of I_m. It can be verified that

$$U_{\alpha,\beta} \in \mathbb{O}^{m\times\ell} \quad \text{for all } \alpha, \beta \in \mathbb{R} \text{ satisfying } \alpha^2 + \beta^2 = 1,$$

and if also $\alpha \neq 0$,

$$\Delta E(U_{\alpha,\beta}, V) = \operatorname{tr}(U_{\alpha,\beta}^\top E(V) U_{\alpha,\beta}) - \operatorname{tr}(M_V) = \alpha^2[\lambda_p(E(V)) - \lambda_{\pi_p}(E(V))] > 0. \tag{4.8}$$

The rest of the proof is to look for some $\alpha, \beta \in \mathbb{R}$ satisfying $\alpha^2 + \beta^2 = 1$ such that $f(U_{\alpha,\beta}) > f(V)$. To this end, we recall (4.1). So, equivalently, we need to make $\Delta f(U_{\alpha,\beta}, V) > 0$. By (4.3) and (4.8), it suffices to find $\alpha, \beta \in \mathbb{R}$, $\alpha^2 + \beta^2 = 1$, $\alpha \neq 0$, and $\delta(U_{\alpha,\beta}, V) = 0$ which is equivalent to either $\phi_B(U_{\alpha,\beta}) = \phi_B(V)$ or $\phi_C(U_{\alpha,\beta}) = \phi_C(V)$.

We have

$$U_{\alpha,\beta}^\top V = V^\top U_{\alpha,\beta} = I_\ell - (1-\beta)\boldsymbol{e}_p\boldsymbol{e}_p^\top, \tag{4.9}$$

$$\begin{aligned} \operatorname{tr}(U_{\alpha,\beta}^\top [I_m - VV^\top] BV) &= \operatorname{tr}(U_{\alpha,\beta}^\top BV) - \phi_B(V) + (1-\beta)\operatorname{tr}(\boldsymbol{e}_p\boldsymbol{e}_p^\top V^\top BV) \\ &= -\operatorname{tr}(\boldsymbol{e}_p\boldsymbol{y}^\top BV) + (1-\beta)\operatorname{tr}(\boldsymbol{e}_p\boldsymbol{e}_p^\top V^\top BV) \\ &= -\boldsymbol{y}^\top BV\boldsymbol{e}_p + (1-\beta)\boldsymbol{e}_p^\top V^\top BV\boldsymbol{e}_p \\ &= -(1-\beta)\boldsymbol{v}_p^\top B\boldsymbol{v}_p + \alpha\boldsymbol{u}^\top B\boldsymbol{v}_p + (1-\beta)\boldsymbol{v}_p^\top B\boldsymbol{v}_p \\ &= \alpha\boldsymbol{u}^\top B\boldsymbol{v}_p, \end{aligned} \tag{4.10}$$

3) Such $\boldsymbol{u}$ can be found even if $\lambda_p = \lambda_{p-1} = \lambda_{\pi_{p-1}}$.

and similarly,

$$\mathrm{tr}(U_{\alpha,\beta}^{\top}[I_m - VV^{\top}]CV) = \alpha \boldsymbol{u}^{\top} C \boldsymbol{v}_p. \tag{4.11}$$

Since V is a global maximizer, we have, by (4.9)–(4.11), and by applying Theorem 2.3 to $U_{\alpha,\beta}$,

$$\mathrm{tr}(U_{\alpha,\beta}^{\top} E(V) U_{\alpha,\beta}) - \mathrm{tr}(M_V) + 4\frac{\alpha^2(\boldsymbol{u}^{\top}B\boldsymbol{v}_p)(\boldsymbol{u}^{\top}C\boldsymbol{v}_p)}{\phi_B(V)} \leqslant 0. \tag{4.12}$$

This inequality, together with (4.8), implies that

$$(\boldsymbol{u}^{\top}B\boldsymbol{v}_p)(\boldsymbol{u}^{\top}C\boldsymbol{v}_p) < 0. \tag{4.13}$$

On the other hand, we have, by (4.7),

$$\begin{aligned}\phi_B(U_{\alpha,\beta}) &= \mathrm{tr}(U_{\alpha,\beta}^{\top} B U_{\alpha,\beta}) \\ &= \phi_B(V) - 2\,\mathrm{tr}(\boldsymbol{e}_p\boldsymbol{y}^{\top}BV) + \mathrm{tr}(\boldsymbol{e}_p\boldsymbol{y}^{\top}B\boldsymbol{y}\boldsymbol{e}_p^{\top}) \\ &= \phi_B(V) - 2\boldsymbol{y}^{\top}B\boldsymbol{v}_p + \boldsymbol{y}^{\top}B\boldsymbol{y} \\ &= \phi_B(V) + \underbrace{\alpha^2\boldsymbol{u}^{\top}B\boldsymbol{u} + 2\alpha\beta\boldsymbol{u}^{\top}B\boldsymbol{v}_p + (\beta^2-1)\boldsymbol{v}_p^{\top}B\boldsymbol{v}_p}_{=:\,h_B(\alpha,\beta)}, \end{aligned} \tag{4.14}$$

$$\phi_C(U_{\alpha,\beta}) = \phi_C(V) + \underbrace{\alpha^2\boldsymbol{u}^{\top}C\boldsymbol{u} + 2\alpha\beta\boldsymbol{u}^{\top}C\boldsymbol{v}_p + (\beta^2-1)\boldsymbol{v}_p^{\top}C\boldsymbol{v}_p}_{=:\,h_C(\alpha,\beta)}. \tag{4.15}$$

Because V is a global maximizer, we have from (4.3) that

$$0 \geqslant f(U_{1,0}) - f(V) = \frac{\phi_B(V)\Delta E(U_{1,0},V) + \delta(U_{1,0},V)}{\phi_B(U_{1,0})}. \tag{4.16}$$

By (4.8), $\Delta E(U_{1,0},V) = \mathrm{tr}(U_{1,0}^{\top}E(V)U_{1,0}) - \mathrm{tr}(M_V) > 0$, and therefore, (4.16) implies

$$\begin{aligned} 0 > \delta(U_{1,0},V) &= [\phi_B(U_{1,0}) - \phi_B(V)][\phi_C(U_{1,0}) - \phi_C(V)] \\ &= h_B(1,0)\,h_C(1,0) \\ &= [\boldsymbol{u}^{\top}B\boldsymbol{u} - \boldsymbol{v}_p^{\top}B\boldsymbol{v}_p][\boldsymbol{u}^{\top}C\boldsymbol{u} - \boldsymbol{v}_p^{\top}C\boldsymbol{v}_p]. \end{aligned} \tag{4.17}$$

There are two possible cases implied by (4.17):

$$\begin{cases} \textbf{Case 1}: \ \boldsymbol{u}^{\top}B\boldsymbol{u} > \boldsymbol{v}_p^{\top}B\boldsymbol{v}_p \quad \text{and} \quad \boldsymbol{u}^{\top}C\boldsymbol{u} < \boldsymbol{v}_p^{\top}C\boldsymbol{v}_p, \\ \textbf{Case 2}: \ \boldsymbol{u}^{\top}B\boldsymbol{u} < \boldsymbol{v}_p^{\top}B\boldsymbol{v}_p \quad \text{and} \quad \boldsymbol{u}^{\top}C\boldsymbol{u} > \boldsymbol{v}_p^{\top}C\boldsymbol{v}_p. \end{cases} \tag{4.18}$$

We will show that for each case, there exist real $\alpha \neq 0$ and real $\beta \neq 0$ satisfying $\alpha^2 + \beta^2 = 1$ and $f(U_{\alpha,\beta}) > f(V)$. The solutions of either $h_B(\alpha,\beta) = 0$ or $h_C(\alpha,\beta) = 0$ turn out to be ones we are looking for.

Suppose Case 1. Consider the equation $h_B(\alpha,\beta) = 0$ which can be regarded as a quadratic equation of α with a parameter β. Since

$$4(\beta\boldsymbol{u}^{\top}B\boldsymbol{v}_p)^2 - 4(\beta^2-1)(\boldsymbol{u}^{\top}B\boldsymbol{u})(\boldsymbol{v}_p^{\top}B\boldsymbol{v}_p) > 0 \quad \text{if } |\beta| < 1,$$

there are two roots

$$\alpha_{B,\pm}(\beta) = \frac{-\beta\boldsymbol{u}^{\top}B\boldsymbol{v}_p \pm \sqrt{(\beta\boldsymbol{u}^{\top}B\boldsymbol{v}_p)^2 - (\beta^2-1)(\boldsymbol{u}^{\top}B\boldsymbol{u})(\boldsymbol{v}_p^{\top}B\boldsymbol{v}_p)}}{\boldsymbol{u}^{\top}B\boldsymbol{u}}. \tag{4.19}$$

We claim that

$$\psi_B(\beta) := \alpha_{B,+}^2(\beta) + \beta^2 - 1 = 0 \tag{4.20}$$

has at least one nonzero root $\tilde{\beta}$ in $(-1,1)$. To prove this claim, we note that

$$\psi_B(0)=\frac{\boldsymbol{v}_p^\top B\boldsymbol{v}_p}{\boldsymbol{u}^\top B\boldsymbol{u}}-1<0 \quad (\text{because Case 1 implies } \boldsymbol{u}^\top B\boldsymbol{u}>\boldsymbol{v}_p^\top B\boldsymbol{v}_p).$$

On the other hand, from (4.13), we know $\boldsymbol{u}^\top B\boldsymbol{v}_p\neq 0$. Now

$$\psi_B(-1)=\frac{2\boldsymbol{u}^\top B\boldsymbol{v}_p}{\boldsymbol{u}^\top B\boldsymbol{u}}>0 \quad \text{if } \boldsymbol{u}^\top B\boldsymbol{v}_p>0,$$
$$\psi_B(+1)=-\frac{2\boldsymbol{u}^\top B\boldsymbol{v}_p}{\boldsymbol{u}^\top B\boldsymbol{u}}>0 \quad \text{if } \boldsymbol{u}^\top B\boldsymbol{v}_p<0.$$

So there exists at least one $\tilde{\beta}$ in either $(-1,0)$ or $(0,1)$, depending on wether $\boldsymbol{u}^\top B\boldsymbol{v}_p>0$ or $\boldsymbol{u}^\top B\boldsymbol{v}_p<0$, such that $\psi_B(\tilde{\beta})=0$.

For Case 2, we consider the equation $h_C(\alpha,\beta)=0$. The same argument, except changing B to [4] C, concludes that there exists at least one $\tilde{\beta}$ in either $(-1,0)$ or $(0,1)$, depending on wether $\boldsymbol{u}^\top C\boldsymbol{v}_p>0$ or $\boldsymbol{u}^\top C\boldsymbol{v}_p<0$, such that $\psi_C(\tilde{\beta})=0$, where $\psi_C(\beta):=\alpha_{C,+}^2(\beta)+\beta^2-1$, and

$$\alpha_{C,+}(\beta)=\frac{-\beta\boldsymbol{u}^\top C\boldsymbol{v}_p+\sqrt{(\beta\boldsymbol{u}^\top C\boldsymbol{v}_p)^2-(\beta^2-1)(\boldsymbol{u}^\top C\boldsymbol{u})(\boldsymbol{v}_p^\top C\boldsymbol{v}_p)}}{\boldsymbol{u}^\top C\boldsymbol{u}}. \tag{4.21}$$

Let $\tilde{\alpha}=\alpha_{B,+}(\tilde{\beta})$ for Case 1 and $\tilde{\alpha}=\alpha_{C,+}(\tilde{\beta})$ for Case 2. This $\tilde{\alpha}\neq 0$ and $\tilde{\alpha}^2+\tilde{\beta}^2=1$. So $\Delta E(U_{\tilde{\alpha},\tilde{\beta}},V)>0$, and either $\phi_B(U_{\tilde{\alpha},\tilde{\beta}})=\phi_B(V)$ for Case 1 or $\phi_C(U_{\tilde{\alpha},\tilde{\beta}})=\phi_C(V)$ for Case 2, i.e., always $\delta(U_{\tilde{\alpha},\tilde{\beta}},V)=0$. By (4.3), $\Delta f(U_{\tilde{\alpha},\tilde{\beta}},V)=f(U_{\tilde{\alpha},\tilde{\beta}})-f(V)>0$, i.e., $f(U_{\tilde{\alpha},\tilde{\beta}})>f(V)$ which is a contradiction since V is a global maximizer. The proof is completed. □

As an interesting application of Theorem 4.1, we can give a characterization of a subset of the global maximizers in some special cases, including $C=\eta B$ for some $\eta>0$.

Theorem 4.3. *Suppose C is positive definite. Let V be a global maximizer to* (1.1) *and $\mathbb{E}_1(E(V))$ be the set of all orthonormal eigenbasis corresponding to the ℓ largest eigenvalues of $E(V)$. If $J\in\mathbb{E}_1(E(V))$ satisfies*

$$\delta(J,V)=[\phi_B(J)-\phi_B(V)][\phi_C(J)-\phi_C(V)]\geqslant 0, \tag{4.22}$$

then J is a global maximizer of (1.1)*, and moreover, $\phi_B(V)=\phi_B(J)$ or $\phi_C(V)=\phi_C(J)$.*

Proof. For any $J\in\mathbb{E}_1(E(V))$, we have $\Delta E(J,V)=0$ by (4.1). Use (4.22) to get

$$f(J)-f(V)=\frac{\phi_B(V)\Delta E(J,V)+\delta(J,V)}{\phi_B(J)}=\frac{\delta(J,V)}{\phi_B(J)}\geqslant 0,$$

which implies J is a global maximizer, and furthermore, $\delta(J,V)=0$ implying either $\phi_B(V)=\phi_B(J)$ or $\phi_C(V)=\phi_C(J)$. □

For the special case $C=\eta B$ $(\eta>0)$, $\delta(J,V)=\eta[\phi_B(V)-\phi_B(J)]^2\geqslant 0$ for any V and J of apt sizes. Therefore, by Theorem 4.3, if V is a global maximizer, then any $J\in\mathbb{E}_1(E(V))$ is also a global maximizer. Moreover, $\phi_B(V)=\phi_B(J)$ and $\phi_C(V)=\phi_C(J)$ which imply $\phi_A(V)=\phi_A(J)$ also.

5 A scheme to make progress at a KKT point

As we mentioned at the beginning of Section 4, C is assumed positive definite in this section, too.

In parallel to the vector version of the problem (1.1) [14, Subsection 3.3], we suggest a strategy to select the starting point for any iterative method that is capable of converging to KKT points monotonically. We argue that, under a mild condition (Assumption (5.2) below), as long as the necessary condition

[4] This is where that C is positive definite is required.

in Theorem 4.1 for a global maximizer is not satisfied, such strategy can be applied to return another approximation with a larger objective value.

Suppose $\bar{V}$ is a KKT point, i.e., $\bar{V}$ satisfies (2.9), but the eigenvalues of

$$M_{\bar{V}} = \bar{V}^{\top} E(\bar{V})\bar{V}$$

do not consist of the ℓ largest eigenvalues of $E(\bar{V})$; so $\bar{V}$ is not a global maximizer, according to Theorem 4.1. Such a $\bar{V}$ could be obtained, for example, by an iterative procedure which may no longer be able to make further progress but to stop.

Owing to the constructive proof of Theorem 4.1 in the previous section, we will explain an idea to break out of the KKT point to a new point $\widetilde{V}_0$ that satisfies $f(\widetilde{V}_0) > f(\bar{V})$. Thus the iterative procedure that got to $\bar{V}$ in the first place can continue at the new point $\widetilde{V}_0$ to the next KKT point.

Let the eigen-decomposition of $M_{\bar{V}}$ be

$$M_{\bar{V}} = U \operatorname{diag}(\lambda_{\pi_1}(E(\bar{V})), \dots, \lambda_{\pi_\ell}(E(\bar{V})))U^{\top}, \tag{5.1a}$$

$$\lambda_{\pi_1}(E(\bar{V})) \geqslant \cdots \geqslant \lambda_{\pi_\ell}(E(\bar{V})). \tag{5.1b}$$

Set $V = \bar{V}U$, and note $E(\bar{V}) = E(V)$. Define p and $\boldsymbol{u}$ by (4.5) and (4.6), respectively.

According to the proof of Theorem 4.1, if

$$0 \leqslant \delta(U_{1,0}, V) = (\boldsymbol{u}^{\top} B\boldsymbol{u} - \boldsymbol{v}_p^{\top} B\boldsymbol{v}_p)(\boldsymbol{u}^{\top} C\boldsymbol{u} - \boldsymbol{v}_p^{\top} C\boldsymbol{v}_p),$$

then $f(U_{1,0}) > f(V) = f(\bar{V})$, where

$$U_{1,0} = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_{p-1}, \boldsymbol{u}, \boldsymbol{v}_{p+1}, \dots, \boldsymbol{v}_\ell].$$

So we may simply take $\widetilde{V}_0 = U_{1,0}$.

Otherwise, suppose $\delta(U_{1,0}, V) < 0$. This is the situation where we need the following assumption[5] that should often be true,

$$\boxed{\textbf{Assumption}\text{: } (\boldsymbol{u}^{\top} B\boldsymbol{v}_p)(\boldsymbol{u}^{\top} C\boldsymbol{v}_p) \neq 0.} \tag{5.2}$$

That $\delta(U_{1,0}, V) < 0$ results in two mutually exclusive cases given in (4.18):

- For Case 1, solve $\psi_B(\tilde{\beta}) := \alpha_{B,+}^2(\tilde{\beta}) + \tilde{\beta}^2 - 1 = 0$ for

$$\begin{cases} \tilde{\beta} \in (\ \ 0, 1), & \text{if } \boldsymbol{u}^{\top} B\boldsymbol{v}_p < 0, \\ \tilde{\beta} \in (-1, 0), & \text{if } \boldsymbol{u}^{\top} B\boldsymbol{v}_p > 0, \end{cases} \tag{5.3}$$

where $\alpha_{B,+}(\beta)$ is given by (4.19);

- For Case 2, solve $\psi_C(\tilde{\beta}) := \alpha_{C,+}^2(\tilde{\beta}) + \tilde{\beta}^2 - 1 = 0$ for

$$\begin{cases} \tilde{\beta} \in (\ \ 0, 1), & \text{if } \boldsymbol{u}^{\top} C\boldsymbol{v}_p < 0, \\ \tilde{\beta} \in (-1, 0), & \text{if } \boldsymbol{u}^{\top} C\boldsymbol{v}_p > 0, \end{cases} \tag{5.4}$$

where $\alpha_{C,+}(\beta)$ is given by (4.21).

Finally, let $\tilde{\alpha} = \alpha_{B,+}(\tilde{\beta})$ for Case 1 and $\tilde{\alpha} = \alpha_{C,+}(\tilde{\beta})$ for Case 2, respectively, and take $\widetilde{V}_0 = U_{\tilde{\alpha},\tilde{\beta}}$, where $U_{\alpha,\beta}$ is defined by (4.7). By the proof of Theorem 4.1,

$$f(\widetilde{V}_0) = f(U_{\tilde{\alpha},\tilde{\beta}}) > f(V) = f(\bar{V}).$$

Conceivably, we may seek $\widetilde{V}_0$ by maximizing $f(U_{\alpha,\beta})$ over $\alpha, \beta \in \mathbb{R}$ subject to $\alpha^2 + \beta^2 = 1$ because a better $\widetilde{V}_0$ in terms of a larger value of the objective function will be gotten than $U_{\tilde{\alpha},\tilde{\beta}}$ we just defined. It turns out that maximizing $f(U_{\alpha,\beta})$ is not trivial and may have to be solved iteratively, too.

The above construction requires (5.2). Can it fail? The following proposition says it cannot if $\bar{V}$ is a local but non-global maximizer.

[5] Recall (4.13) in the proof of Theorem 4.1, where although $(\boldsymbol{u}^{\top} B\boldsymbol{v}_p)(\boldsymbol{u}^{\top} C\boldsymbol{v}_p) < 0$, only its implication that both $\boldsymbol{u}^{\top} B\boldsymbol{v}_p$ and $\boldsymbol{u}^{\top} C\boldsymbol{v}_p$ are nonzero were used later in the proof.

Algorithm 5.1. Breaking out of a KKT point

Given A KKT point $\bar{V}$ which is not an orthonormal eigenbasis corresponding to the ℓ largest eigenvalues of $E(\bar{V})$, this algorithm computes $\widetilde{V}_0 \in \mathbb{O}^{m\times\ell}$ such that $f(\widetilde{V}_0) \geqslant f(\bar{V})$.

compute the eigen-decomposition of $M_{\bar{V}}$ as in (5.1), and let $V = \bar{V}U$;
compute p and $\boldsymbol{u}$ by (4.5) and (4.6), respectively;
while $(\boldsymbol{u}^\top B\boldsymbol{v}_p)(\boldsymbol{u}^\top C\boldsymbol{v}_p) = 0$ and $p \leqslant \ell$ **do**
    $p := p + 1$;
end while
if $p \leqslant \ell$ **then**
    if $0 \leqslant \delta(U_{1,0}, V)$ **then**
        $\widetilde{V}_0 = U_{1,0}$;
    else
        solve (5.3) in **Case 1** or (5.4) in **Case 2** for $\tilde{\beta}$;
        $\tilde{\alpha} = \alpha_{B,+}(\tilde{\beta})$ if **Case 1** or $\tilde{\alpha} = \alpha_{C,+}(\tilde{\beta})$ if **Case 2**;
        $\widetilde{V}_0 = U_{\tilde{\alpha},\tilde{\beta}}$;
    end if
else
    $\widetilde{V}_0 = \bar{V}$; `/* no improvement is made */`
end if

Proposition 5.1. *Suppose C is positive definite, and let $\bar{V}$ be a local but non-global maximizer $\bar{V}$ of* (1.1). *Let the eigen-decomposition of $M_{\bar{V}}$ be given by* (5.1), *and set $V = \bar{V}U = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell]$,*

$$p := \min_{1\leqslant i\leqslant \ell}\{i \,:\, \lambda_{\pi_i}(E(\bar{V})) < \lambda_i(E(\bar{V}))\},$$
$$E(\bar{V})\boldsymbol{u} = \lambda_p(E(\bar{V}))\boldsymbol{u} \quad \text{and} \quad \bar{V}^\top \boldsymbol{u} = \boldsymbol{0}.$$

Then $(\boldsymbol{u}^\top B\boldsymbol{v}_p)(\boldsymbol{u}^\top C\boldsymbol{v}_p) < 0$. As a result, (5.2) *holds.*

Proof. By Theorem 4.1, we know that $1 \leqslant p \leqslant \ell$. For any two $\alpha, \beta \in \mathbb{R}$ with $\alpha^2 + \beta^2 = 1$, let $U_{\alpha,\beta} \in \mathbb{O}^{m\times\ell}$ be as defined in (4.7). Since V is also a local but non-global maximizer, by (4.9)–(4.11) and by applying Theorem 2.3 to $U_{\alpha,\beta}$, we have the inequality (4.12). By (4.8), $\mathrm{tr}(U_{\alpha,\beta}^\top E(V)U_{\alpha,\beta}) - \mathrm{tr}(M_V) > 0$ for $0 < |\alpha| < 1$. Now use (4.12) to conclude $(\boldsymbol{u}^\top B\boldsymbol{v}_p)(\boldsymbol{u}^\top C\boldsymbol{v}_p) < 0$. □

In general for a KKT point that is not a local maximizer, we do not know if the assumption (5.2) fails or not. But we propose the following workaround. If $(\boldsymbol{u}^\top B\boldsymbol{v}_p)(\boldsymbol{u}^\top C\boldsymbol{v}_p) = 0$ and $p < \ell$, we can simply let $p := p + 1$ and check if $(\boldsymbol{u}^\top B\boldsymbol{v}_p)(\boldsymbol{u}^\top C\boldsymbol{v}_p)$ is nonzero. This process repeats until $(\boldsymbol{u}^\top B\boldsymbol{v}_p)(\boldsymbol{u}^\top C\boldsymbol{v}_p) \neq 0$ is encountered, or $p = \ell$ but still $(\boldsymbol{u}^\top B\boldsymbol{v}_p)(\boldsymbol{u}^\top C\boldsymbol{v}_p) = 0$, the worst case for which the workaround fails, too. But that should be rare. The complete algorithm, including this workaround, is summarized in Algorithm 5.1.

The above strategy requires solving either $\psi_B(\beta) = 0$ or $\psi_C(\beta) = 0$. Both equations can be turned into quadratic equations in β^2 of the form:

$$\alpha_4\beta^4 + \alpha_2\beta^2 + \alpha_0 = 0, \tag{5.5}$$

where for $\psi_X(\beta) = 0$ with $X = B$ or C, the coefficients of (5.5) are given by

$$\begin{aligned}
&x_{11} = \boldsymbol{u}^\top X\boldsymbol{u}, \quad x_{22} = \boldsymbol{v}_p^\top X\boldsymbol{v}_p, \quad x_{12} = \boldsymbol{u}^\top X\boldsymbol{v}_p,\\
&\eta = x_{12}^2 - x_{11}x_{22},\\
&\alpha_4 = 4x_{12}\eta - (x_{11}^2 + x_{12}^2 + \eta)^2,\\
&\alpha_2 = 4x_{11}x_{12}x_{22} + 2x_{11}(x_{11} - x_{22})(x_{11}^2 + x_{12}^2 + \eta),
\end{aligned}$$

$$\alpha_0 = -x_{11}^2(x_{11} - x_{22})^2.$$

Depending on the cases, one can solve (5.5) for the right β. Note

$$\alpha_2^2 - 4\alpha_0\alpha_4 = 16x_{11}^2x_{12}^2[x_{12}(x_{11}^2 - x_{22}^2) + x_{22}^2].$$

Alternatively, one can certainly employ some iterative methods such as the bisection or the Newton iteration to solve $\psi_X(\beta) = 0$ for β.

6 Concluding remarks

Analogously to that maximizing $\mathrm{tr}(V^\top CV)$ on $\mathbb{O}^{m\times\ell}$ is a subspace version of $\max_{\|\boldsymbol{v}\|_2=1}\boldsymbol{v}^\top C\boldsymbol{v}$ whose solution is the dominant eigenpair of C, what we have studied here

$$\max_{V^\top V=I_\ell} f(V) = \max_{V^\top V=I_\ell}\left\{\frac{\mathrm{tr}(V^\top AV)}{\mathrm{tr}(V^\top BV)} + \mathrm{tr}(V^\top CV)\right\}, \tag{1.1}$$

is a subspace version of optimizing the sum of two Rayleigh quotients on the unit sphere. We have made a number of discoveries, including the connection between (1.1) and the nonlinear linear eigenvalue problem (2.9), and the elegant necessary characterization of a global maximizer in terms of the extreme eigenvalues of (2.9) that leads to the SCF iteration for numerical solving (1.1) in [16]. We also explain how to march out of a KKT point that is not a global maximizer.

Although (2.9) is about maximization, $\min f(V)$, if needed, can be turned into one of (1.1) (with a different f). In fact,

$$\min f(V) = -\max[-f(V)] = -\max\left\{\frac{\mathrm{tr}(V^\top[-A]V)}{\mathrm{tr}(V^\top BV)} + \mathrm{tr}(V^\top[-C]V)\right\}.$$

We mentioned that (1.1) is a subspace version of optimizing the sum of two Rayleigh quotients on the unit sphere (the case $\ell = 1$). But (1.1) is only one kind among variously conceivable subspace versions. Other versions may deserve attentions, too. For example, it is easy to see that when $\ell = 1$, (1.1) is the same as

$$\max_{V^\top V=I_\ell}\{\mathrm{tr}((V^\top AV)(V^\top BV)^{-1}) + \mathrm{tr}(V^\top CV)\}. \tag{6.1}$$

Thus (6.1) can be considered as another subspace version. We believe similar lines of arguments of ours so far can be used to investigate (6.1).

It is interesting to note that (6.1) is essentially equivalent to

$$\max_{\mathrm{rank}(Y)=\ell}\{\mathrm{tr}((Y^\top\widetilde{A}Y)(Y^\top\widetilde{B}Y)^{-1}) + \mathrm{tr}((Y^\top\widetilde{C}Y)(Y^\top\widetilde{D}Y)^{-1})\}, \tag{6.2}$$

where $\widetilde{A}$ and $\widetilde{C}$ are symmetric, and $\widetilde{B}$ and $\widetilde{D}$ are symmetric and positive definite. In fact, we can set

$$Z := \widetilde{D}^{1/2}Y, \quad A := \widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}, \quad B := \widetilde{D}^{-1/2}\widetilde{B}\widetilde{D}^{-1/2}, \quad C := \widetilde{D}^{-1/2}\widetilde{C}\widetilde{D}^{-1/2}$$

to translate (6.2) into

$$\max_{\mathrm{rank}(Z)=\ell}\{\mathrm{tr}((Z^\top AZ)(Z^\top BZ)^{-1}) + \mathrm{tr}((Z^\top CZ)(Z^\top Z)^{-1})\}. \tag{6.3}$$

Now substitute $Z = VR$ (the thin QR factorization of Z) into (6.3) to arrive at (6.1).

Acknowledgements The first author was supported by National Natural Science Foundation of China (Grant Nos. 11101257 and 11371102), and the Basic Academic Discipline Program, the 11th Five Year Plan of 211 Project for Shanghai University of Finance and Economics. Part of this work was done while the first author was a visiting scholar at the Department of Mathematics, University of Texas at Arlington from February 2013 to January 2014. The second author was supported by National Science Foundation of USA (Grant Nos. 1115834 and 1317330), and a Research Gift Grant from Intel Corporation. The authors are grateful to two anonymous referees for their careful reading and helpful comments and suggestions.

References

1 Absil P A, Mahony R, Sepulchre R. Optimization Algorithms On Matrix Manifolds. Princeton, NJ: Princeton University Press, 2008

2 Chu M T, Trendafilov N T. The orthogonally constrained regression revisited. J Comput Graph Stat, 2001, 10: 746–771

3 Dundar M M, Fung G, Bi J, et al. Sparse Fisher discriminant analysis for computer aided detection. In: Proceedings of the SIAM International Conference on Data Mining, 2005

4 Edelman A, Arias T A, Smith S T. The geometry of algorithms with orthogonality constraints. SIAM J Matrix Anal Appl, 1999, 20: 303–353

5 Fung E, Ng M K. On sparse Fisher discriminant method for microarray data analysis. Bioinformation, 2007, 2: 230–234

6 Helmke U, Moore J B. Optimization and Dynamical Systems. London: Springer-Verlag, 1994

7 Ng M K, Liao, L Z, Zhang L H. On sparse linear discriminant analysis algorithm for high-dimensional data classification. Numer Linear Algebra Appl, 2011, 18: 223–235

8 Nocedal J, Wright S. Numerical Optimization, 2nd ed. New Yrok: Springer, 2006

9 Primolevo G, Simeone O, Spagnolini U. Towards a joint optimization of scheduling and beamforning for MIMO downlink. In: 2006 IEEE Ninth International Symposium on Spread Spectrum Techniques and Applications, 2006, 493–497

10 Shen H, Diepold K, Hüper K. A geometric revisit to the trace quotient problem. In: Proceedings of the 19th International Symposium of Mathematical Theory of Networks and Systems (MTNS 2010), 2010

11 Stewart G W. Matrix Algorithms, vol. II: Eigensystems. Philadelphia: SIAM, 2001

12 Wen Z, Yin W. A feasible method for optimization with orthogonality constraints. Math Programming, 2013, 142: 397–434

13 Yang W H, Zhang L H, Song R Y. Optimality conditions of the nonlinear programming on Riemannian manifolds. Pacific J Opt, 2013, in press

14 Zhang L H. On optimizing the sum of the Rayleigh quotient and the generalized Rayleigh quotient on the unit sphere. Comput Opt Appl, 2013, 54: 111–139

15 Zhang L H, Li R C. Maximization of the sum of the trace ratio on the Stiefel manifold. Technical Report 2013-04, Department of Mathematics, University of Texas at Arlington (2013). Available at http://www.uta.edu/math/preprint/

16 Zhang L H, Li R C. Maximization of the sum of the trace ratio on the Stiefel manifold, II: Computation. Sci China Math, 2015, doi: 10.1007/s11425-014-4825-z, in press

17 Zhang L H, Yang W H, Liao L Z. A note on the trace quotient problem. J Optim Lett, 2013, doi: 10.1007/s11590-013-0680-z, in press