Math Meth Oper Res (2017) 85:453–490
DOI 10.1007/s00186-017-0579-z

ORIGINAL ARTICLE

A space decomposition scheme for maximum eigenvalue functions and its applications

Ming Huang[1,2] · **Yue Lu**[3] · **Li Ping Pang**[4] · **Zun Quan Xia**[4]

Received: 8 December 2015 / Published online: 18 March 2017

Abstract In this paper, we study nonlinear optimization problems involving eigenvalues of symmetric matrices. One of the difficulties in solving these problems is that the eigenvalue functions are not differentiable when the multiplicity of the function is not one. We apply the $\mathcal{U}$-Lagrangian theory to analyze the largest eigenvalue function of a convex matrix-valued mapping which extends the corresponding results for linear mapping in the literature. We also provides the formula of first-and second-order derivatives of the $\mathcal{U}$-Lagrangian under mild assumptions. These theoretical results provide us new second-order information about the largest eigenvalue function along a suitable smooth manifold, and leads to a new algorithmic framework for analyzing the underlying optimization problem.

Keywords Nonsmooth optimization · Eigenvalue optimization · Matrix-convex · Semidefinite programming · $\mathcal{VU}$-decomposition · $\mathcal{U}$-Lagrangian · Smooth manifold · Second-order derivative · Bilinear matrix inequality

Mathematics Subject Classification 15A18 · 52A41 · 65K10 · 90C25 · 90C22 · 49J52

✉ Ming Huang
huangming0224@163.com

1 School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China

2 Department of Mathematics, Dalian Maritime University, Dalian 116026, China

3 School of Mathematical Sciences, Tianjin Normal University, Tianjin 300387, China

4 CORA, School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

1 Introduction

Recently there have been significant advances in the analysis of eigenvalue problems. As mentioned in Lewis and Overton (1996), eigenvalue optimization problems became an independent area of research with both theoretical and practical aspects in the 1980s. Early contributions owe to Cullum et al. (1975), Lewis and Overton (1996), Overton (1992). Eigenvalue optimization has a wide spectrum of applications in physics, engineering, statistics, and finance. This spectrum includes composite materials (Cox and Lipton 1996), quantum computational chemistry (Zhao et al. 2004), optimal system design (Apkarian et al. 2004; Noll et al. 2004), shape optimization (Díaz and Kikuchi 1992), pole placement in linear system theory, robotics, relaxations of combinatorial optimization problems (Helmberg et al. 2000), experimental design (Vandenberghe and Boyd 1996) etc. Optimization problems involving eigenvalues of symmetric matrices arise in many applications, see e.g., Lewis and Overton (1996) and references therein. One of the main difficulties in analyzing such problems is that the eigenvalues, as functions of a symmetric matrix, are not differentiable at optimal points whose multiplicities are greater than one. Typical smooth optimization techniques could not be applied directly because of its nonsmooth property.

Nonsmooth analysis of eigenvalues plays an essential role in designing efficient algorithms for eigenvalue optimization (Lewis 1996). In the 1970s and 1980s first-order algorithms for optimization of nonsmooth functions were developed and applied to eigenvalue optimization problems. At the same time various attempts were made to develop second-order theory for nonsmooth optimization problems. Much pioneering work has been contributed by Overton (1992) beginning in the 1980's, and developed further in Noll and Apkarian (2005), Overton and Womersley (1993, 1995), where Newton-type methods are considered. Further to be mentioned among the earliest contributions are Cullum et al. (1975), Fletcher (1985) and Shapiro and Fan (1995).

In this paper we will study the following model

$$(P) \quad \min_{x \in R^m} f(x) := \lambda_1(A(x)),$$

where the mapping $A : R^m \to S_n$, $\lambda_1(\cdot)$ is the maximum eigenvalue function. The problem (P) was proposed in 1970s by Cullum et al. (1975). Overton and Womersley, Hiriart-Urruty and Ye in 1995 respectively presented its optimality conditions and sensitivity analysis, see Overton and Womersley (1993) and Hiriart-Urruty and Ye (1995). In this paper, we focus on the second-order theory of (P), where the inner mapping A is not necessarily affine about x. We called the problem (P) *a general problem*. The difficulty of the problem is that the objective function is probably not smooth at the optimal solutions. Actually, it is well known that functions depending on eigenvalues are not continuously differentiable at points where eigenvalues are not simple. What is worse is that the optimization process generally leads to points where the eigenvalues coalesce, triggering the non-differentiability of the objective function. A theoretical explanation of this phenomenon can be found in Pataki (1998), in truth for a more general eigenvalue function. In this work, we assume that the multiplicity r of $\lambda_1(A(x^*))$ at an optimal point x^* is known, then the problem is recast as minimizing the maximum eigenvalue subject to the constraint that its multiplicity is constant, i.e., the matrices locate on certain smooth manifold: the set $\mathcal{M}_r$ of matrices whose largest eigenvalue has multiplicity r. We adopt the technique of local C^2-parametrization of (P) to develop a successive quadratic programming method in this paper. Motivated by the work of Lemaréchal et al. (2000), where the authors present the $\mathcal{VU}$-decomposition theory and show that f appears to be smooth on the $\mathcal{U}$-subspace and may have some kind of related Hessian, and the nonsmoothness of f is concentrated essentially on $\mathcal{V}$-subspace, we apply the $\mathcal{U}$-Lagrangian theory to the more general case: the function of the largest eigenvalues with the *convex matrix mapping*, which is defined in Sect. 2.

For the problem (P), the convex eigenvalue optimization has a wide range of research. In particular, standard semidefinite programming (SDP) can be transformed into an eigenvalue optimization problem (Helmberg and Rendl 2000; Oustry 2000). The role of convex analysis was first emphasized by Bellman and Fan in Bellman and Fan (1963), and this point of view was developed further in Hiriart-Urruty and Lemaréchal (1993), Lewis and Overton (1996) more recently. The approach to solve convex eigenvalue optimization can be classified into two types, the interior-point methods and the nonsmooth optimization methods. The first interior-point methods for solving convex eigenvalue optimization were presented by Nesterov and Nemirovskii (1988). In addition, most of the interior-point schemes proposed in the early 1990s were path-following or potential reduction methods (Alizadeh et al. 1998; Kojima et al. 1998; Nesterov 1997). However, when the problem is defined over large matrix variables or a huge number of constraints, interior point methods usually grow terribly slow and consume huge amounts of memory. Helmberg contributed a series of papers (Helmberg and Rendl 2000 with Rendl, and Helmberg and Kiwiel 2002 with Kiwiel), where in particular spectral bundle methods are discussed. By using $\mathcal{U}$-Lagrangian theory, we present a second-order method, which is different from the above methods. Our approach is more elementary and direct, and emphasizes a pleasing parallel between local and global properties.

The goal of this work is to construct a general framework of second-order analysis. Our contribution is to show that second-order $\mathcal{VU}$ decomposition results can be obtained making use of the idea from Oustry (1999), where the inner mapping is required to be affine. However, we can obtain $\mathcal{VU}$ space decomposition of the largest eigenvalues with the convex matrix inner mapping, and we don't make the assumption that it is an affine mapping, i.e., our results are applicable to a broad problem class, in which nonlinear models are permitted. Another major difference is that here the linear operator $DA(\cdot)$ relies on some point, but (Oustry 1999) reduces to $\mathcal{A}$, now attached to space decomposition, which will lead to our conclusions being more general and complicated. Meanwhile, we assume the regular condition holds, under which we can use the vectors of the $\mathcal{V}$-space to generate an implicit function therein from which a smooth trajectory tangent to $\mathcal{U}$ can be defined. This condition plays a role similar to that of constraint qualification conditions in nonlinear programming. Once the regular condition is satisfied, the maximum eigenvalue function λ_1 has a second-order expansion along the associated trajectory.

These theoretical results naturally lead to $\mathcal{VU}$-decomposition algorithms. The resulting algorithms iteratively make a step in the $\mathcal{V}$-subspace, followed by a $\mathcal{U}$-Newton step in order to obtain superlinear convergence. Moreover, it is applied to a class of important optimization programs, i.e., the bilinear matrix inequality problem. The maximum eigenvalue function with matrix variable is also researched. In particular, we present a weaker condition than transversality and prove these results still hold.

The rest of the paper is organized as follows. In Sect. 2, we give some preliminaries including a brief introduction about $\mathcal{U}$-Lagrangian theory and differential geometry concepts. The main results are given in Sect. 3. It is shown that $\mathcal{U}$-Lagrangian of the maximum eigenvalues is a C^∞ convex function, which identifies locally the ridge of the maximum eigenvalue function λ_1. The second-order expansion of λ_1 is derived and the second-order derivatives are explicitly computed. In Sect. 4 we present the second-order expansion of the maximum eigenvalue function λ_1 along some smooth trajectory tangent to $\mathcal{U}$-space; in addition, we describe an algorithm framework, and analyze its local superlinear convergence rate. In Sect. 5, the algorithm is applied to solve the bilinear matrix inequality problem, indicating its potential for application. Finally, we conclude this work by pointing out some promising research issues.

Next we introduce the main notations and terminologies used in the remaining sections. S_n is the set of $n \times n$ symmetric matrices, S_n^+ stands for the set of $n \times n$ positive semidefinite symmetric matrices, and $A \succ 0$ indicates that the matrix A is positive definite. $\mathrm{proj}_{\mathcal{U}} : R^m \to \mathcal{U}$ is a projection operator that projects a vector from R^m onto the subspace $\mathcal{U}$, $\mathrm{proj}^*_{\mathcal{U}} : \mathcal{U} \mapsto R^m$ is the canonical injection $\mathcal{U} \ni u \mapsto u \oplus 0 \in R^m$. $A \cdot B := \mathrm{tr} AB$ denotes the Fröbenius scalar product of $A, B \in S_n$, and $A^\dagger$ indicates the Moore-Penrose inverse of A, $\mathcal{A}^\star : S_n \to R^m$ is the adjoint operator of the linear operator $\mathcal{A} : R^m \to S_n$. Let $r \geq 1$ be the multiplicity of the largest eigenvalue $\lambda_1(A)$ of A , i.e., A lies on the submanifold $\mathcal{M}_r := \{A \in S_n : \lambda_1(A) = \cdots = \lambda_r(A) > \lambda_{r+1}(A)\}$, where $\mathcal{M}_r$ is a C^∞-submanifold of S_n. Let $E_1(A)$ be the first eigenspace associated with λ_1, $Q_1(A)$ be an orthonormal basis of $E_1(A)$, $T_{\mathcal{M}}(A)$ and $N_{\mathcal{M}}(A)$ be the tangent and normal spaces to the submanifold $\mathcal{M}$ at $A \in \mathcal{M}$, respectively. $\mathrm{rank}(A)$ denotes the rank of the matrix A. The notation $DA(x)$ is used for the differential of the mapping $A(\cdot)$ at x, i.e., $DA(x)$ is a linear mapping from R^m into S_n defined by $DA(x)y = \sum_{i=1}^m y_i A_i(x)$, where $A_i(x) = \partial A(x)/\partial x_i$ are $n \times n$ partial derivatives. For the other signs, we refer to Hiriart-Urruty and Lemaréchal (1993), Rockafellar (1970).

2 Preparation and preliminary results

2.1 Some preliminaries about $\mathcal{VU}$-decomposition

In this subsection, we first review some background materials on $\mathcal{VU}$-space decomposition and $\mathcal{U}$-Lagrangian theory in a general framework used in this paper.

For a convex function f at a given point $\bar{x} \in R^m$ where f is finite, let g be any subgradient in $\partial f(\bar{x})$. Then, letting $\mathrm{lin} Y$ denote the linear hull of a given set Y, the smallest linear subspace containing Y. The orthogonal subspaces

$$\mathcal{V}(\bar{x}) := \text{lin}(\partial f(\bar{x}) - g) \quad \textit{and} \quad \mathcal{U}(\bar{x}) := \mathcal{V}(\bar{x})^{\perp} \tag{2.1}$$

define the $\mathcal{VU}$-space decomposition at $\bar{x}$ of Lemaréchal et al. (2000), i.e., $\mathcal{V}(\bar{x})$ and $\mathcal{U}(\bar{x})$ are respectively the subspaces parallel and orthogonal affine hull of the set $\partial f(\bar{x})$. These spaces represent the directions from $\bar{x}$ for which f behaves nonsmoothly ($\mathcal{V}(\bar{x})$) and smoothly ($\mathcal{U}(\bar{x})$). The goal is then to find a smooth function that describes f in the directions of $\mathcal{U}(\bar{x})$. We use the compact notation $\oplus$ for such decomposition, and write $R^m = \mathcal{U}(\bar{x}) \oplus \mathcal{V}(\bar{x})$. Similarly, we denote

$$R^m \ni \bar{x} = \bar{x}_{\mathcal{U}} \oplus \bar{x}_{\mathcal{V}} \in \mathcal{U}(\bar{x}) \times \mathcal{V}(\bar{x}).$$

From (2.1), the relative interior of $\partial f(\bar{x})$, denoted by $\text{ri}\partial f(\bar{x})$, is the interior of $\partial f(\bar{x})$ relative to its affine hull, a manifold that is parallel to $\mathcal{V}(\bar{x})$. Accordingly,

$$\bar{g} \in \text{ri}\partial f(\bar{x}) \quad \Rightarrow \quad \bar{g} + (B(0,\eta) \cap \mathcal{V}(\bar{x})) \subset \partial f(\bar{x}) \quad \text{for some } \eta > 0,$$

where $B(0, \eta)$ denotes a ball in R^m centered at 0 with radius η.

Now we list the following two equivalent definition forms used in the subsequent discussions, which come from the literature (Lemaréchal et al. 2000).

Proposition 2.1 *For a finite-valued convex function f, $\bar{x} \in R^m$ and $g \in \partial f(\bar{x})$ are given, we have*

(i) $\mathcal{U}(\bar{x})$ is the subspace where $f'(\bar{x}; \cdot)$ is linear. In other words, $\mathcal{U}(\bar{x})$ is the subspace where $\phi(t) := f(x + td)$ is differentiable at 0, i.e.,

$$\mathcal{U}(\bar{x}) = \{d \in R^m : f'(\bar{x}; d) = -f'(\bar{x}; -d)\},$$

and $\mathcal{V}(\bar{x}) = \mathcal{U}(\bar{x})^{\perp}$; where $f'(x; d)$ is the directional derivative of the function f at the point x along the direction d.

(ii) For any $g \in \text{ri}\partial f(\bar{x})$, $\mathcal{U}(\bar{x})$ and $\mathcal{V}(\bar{x})$ are, respectively, the normal and tangent cones to $\partial f(\bar{x})$ at g, denoted by $T_{\partial f(\bar{x})}(g)$ and $N_{\partial f(\bar{x})}(g)$.

Based on the above proposition about $\mathcal{VU}$-decomposition, we now present some definitions and properties of the $\mathcal{U}$-Lagrangian for a convex function f. Given a subgradient $\bar{g} \in \partial f(\bar{x})$ with $\mathcal{V}$-component $\bar{g}_{\mathcal{V}}$, the $\mathcal{U}$-Lagrangian of f at the primal-dual pair $(\bar{x}, \bar{g})$, depending on $\bar{g}_{\mathcal{V}}$, is defined by

$$\mathcal{U}(\bar{x}) \ni u \mapsto L_{\mathcal{U}}(u) = \min_{v \in \mathcal{V}(\bar{x})} \{f(\bar{x} + u \oplus v) - \langle \bar{g}_{\mathcal{V}}, v \rangle_{\mathcal{V}}\}, \tag{2.2}$$

where $\langle \cdot, \cdot \rangle_{\mathcal{V}}$ denotes a scalar product induced in the subspace $\mathcal{V}$. When the minimum in (2.2) is attained, the set of corresponding $\mathcal{V}$-space minimizers is defined by

$$w(u) = \{v \in \mathcal{V}(\bar{x}) : L_{\mathcal{U}}(u) = f(\bar{x} + u \oplus v) - \langle \bar{g}_{\mathcal{V}}, v \rangle_{\mathcal{V}}\}. \tag{2.3}$$

When $w(u)$ is nonempty, the associated $\mathcal{U}$-Lagrangian is a convex function that is differentiable at $u = 0$, with

$$\nabla L_{\mathcal{U}}(0) = \bar{g}_{\mathcal{U}} = g_{\mathcal{U}} \quad for\ all\ g \in \partial f(\bar{x}).$$

We summarize some properties of $L_{\mathcal{U}}$ in the following theorems, which will be used in later sections. For details, see reference Lemaréchal et al. (2000) and Oustry (1999).

Theorem 2.1 *The function $L_{\mathcal{U}}$ is well-defined and convex. In addition, if $g \in \mathrm{ri}\partial f(\bar{x})$, the set $w(u)$ is nonempty and the following properties hold:*

1. *The subdifferential of $L_{\mathcal{U}}(u)$ is*
$$\partial L_{\mathcal{U}}(u) = \mathrm{proj}_{\mathcal{U}(\bar{x})}[\partial f(\bar{x} + u \oplus v) \cap (g + \mathcal{U}(\bar{x}))], \tag{2.4}$$
where v is taken arbitrary in $w(u)$.
2. *When $u = 0$, we have $w(0) = \{0\}$ and $L_{\mathcal{U}}(0) = f(\bar{x})$. Moreover, $L_{\mathcal{U}}$ is differentiable at 0 and*
$$\nabla L_{\mathcal{U}}(0) = \mathrm{proj}_{\mathcal{U}(\bar{x})} g. \tag{2.5}$$
3. *The multifunction $u \mapsto \partial L_{\mathcal{U}}(u)$ is continuous at $u = 0$:*
$$\lim_{u \to 0} \partial L_{\mathcal{U}}(u) = \{\nabla L_{\mathcal{U}}(0)\}. \tag{2.6}$$
4. *For all $u \in \mathcal{U}(\bar{x})$, we have*
$$\partial f(\bar{x} + u \oplus v) \cap (g + \mathcal{U}(\bar{x})) = \partial L_{\mathcal{U}}(u) \oplus \{\mathrm{proj}_{\mathcal{V}(\bar{x})} g\} \quad \text{for all } v \in w(u). \tag{2.7}$$
5. *Denoting by $\partial f(\bar{x} + u \oplus w(u)) \cap (g + \mathcal{U}(\bar{x}))$ the right-hand side of (2.6), the multifunction $u \mapsto \partial f(\bar{x} + u \oplus w(u)) \cap (g + \mathcal{U}(\bar{x}))$ is continuous at 0:*
$$\lim_{u \to 0} \partial f(\bar{x} + u \oplus w(u)) \cap (g + \mathcal{U}(\bar{x})) = \{g\}. \tag{2.8}$$
6. *For all $u \in \mathcal{U}(\bar{x})$, $w(u)$ is a nonempty compact convex set that satisfies*
$$\sup_{v \in w(u)} \|v\| = o(\|u\|), \tag{2.9}$$
and the multifunction $u \to w(u)$ is continuous at $u = 0$:
$$\lim_{u \to 0} w(u) = \{0\}.$$

Theorem 2.2 *(1) Given a point $A^* \in S_n$ and a positive number δ_0. Let $\Phi(A) = 0$ be a local equation of the manifold $\mathcal{M}$ in $B(A^*, \delta_0)$, i.e., for each $A \in \Phi^{-1}(0) := \{\Omega \in$*

$B(A^*, \delta_0) : \Phi(\Omega) = 0\}$, $D\Phi(A)$ is surjective, and $\mathcal{M} \cap B(A^*, \delta_0) = \Phi^{-1}(0)$. Then, there exists a scalar δ such that $0 < \delta \leq \delta_0$ and a unique mapping

$$v : T_{\mathcal{M}}(A^*) \cap B(0, \delta) \to N_{\mathcal{M}}(A^*)$$

such that, for all $(u, v) \in (T_{\mathcal{M}}(A^*), N_{\mathcal{M}}(A^*))$,

$$v = v(u).$$

The mapping v is C^∞, and at $u = 0$ we have

$$Dv(0) = 0. \tag{2.10}$$

(2) Let $A^* \in \mathcal{M}$; then there exists $\delta > 0$ such that

$$proj_{N_{\mathcal{M}}(A^*)} d = v(\mathrm{proj}_{N_{\mathcal{M}}(A^*)} d) \tag{2.11}$$

for all $d \in B(0, \delta)$ satisfying $A^* + d \in \mathcal{M}$.

2.2 Matrix convexity

Next we present the definition of matrix convexity and its properties, one can refer to Bonnans and Shapiro (2000) and the reference therein for more proofs and details.

The partial order in the space S_n with respect to the cone S_n^+ is called the *Löwner partial order*. That is, for $A, B \in S_n$, $A \succeq B$ if and only if $A - B$ is a positive semidefinite matrix.

Definition 2.1 (Bonnans and Shapiro 2000) We say that the mapping $G : R^m \to S_n$ is *matrix convex*(on the convex set Q) if it is convex with respect to the Löwner partial order. This means that for any $t \in [0, 1]$ and any $x_1, x_2 \in R^m$ (any $x_1, x_2 \in Q$) the following inequality holds

$$tG(x_1) + (1 - t)G(x_2) \succeq G(tx_1 + (1 - t)x_2).$$

Note that matrix convex is also called positive semidefinite convex (psd-convex).

Proposition 2.2 (Bonnans and Shapiro 2000) *The following results for the mapping $G : R^m \to S_n$ hold.*

(i) $G(x)$ is matrix convex if and only if the real valued function $\psi(x) := z^\top G(x)z$ is convex for any $z \in R^n$.
(ii) If every element $g_{ij}(x), i, j = 1, \cdots, n$, of $G(x)$ is a twice continuously differentiable function, then $G(x)$ is matrix convex if and only if the $m \times m$ matrix $\sum_{i,j=1}^{n} z_i z_j \nabla^2 g_{ij}(x)$ is positive semidefinite for any $(z_1, \ldots, z_n) \in R^n$ and any $x \in R^m$.

(iii) If $G(x)$ is twice continuously differentiable and the block matrix $K(x) := [\nabla^2 g_{ij}(x)]_{i,j=1}^n$ with the dimension $mn \times mn$ is positive semidefinite for any $x \in R^m$, then $G(x)$ is matrix convex.

(iv) If $G(x)$ is matrix convex, then its largest eigenvalue function $\phi(x) := \lambda_1(G(x))$ is convex.

Clearly, any affine mapping $G(x) := A_0 + \sum_{i=1}^m x_i A_i$ is matrix convex, where $A_i, i = 0, 1, \dots, m$ are the given matrices. Based on the above definition and properties, we list some examples of the matrix convex mappings, which can be found in Shapiro (1997).

Example 2.1 Consider the quadratic mapping

$$G(x) := A_0 + \sum_{i=1}^m x_i A_i + \sum_{i,j=1}^m x_i x_j B_{i,j}, \tag{2.12}$$

where $A_i, B_{i,j} \in S_n$ are the given matrices. This mapping is matrix convex if and only if $\sum_{i,j=1}^m x_i x_j z^\top B_{i,j} z \ge 0$ for any $x \in R^m$ and $z \in R^p$. A sufficient condition for that to hold is that the $mn \times mn$ block matrix $[B_{i,j}]_{i,j=1}^m$ is positive semidefinite. Note that, positive semidefiniteness of this block matrix is a sufficient but not necessary condition for matrix convexity of the corresponding quadratic mapping $G(x)$.

The above example arises in many applications in automatic control, finance and design engineering. In particular, solving bilinear matrix inequalities (BMIs) with matrix convex mapping is a prominent application, which may be addressed via eigenvalue optimization.

Example 2.2 Consider the mapping $G(Z) := Z^2$. This mapping is matrix convex. By Proposition 2.2(i), in order to prove this it suffices to show that for any $A, B \in S_n$ and any $z \in R^n$, the real valued function $\psi(t) := z^\top G(A + tB)z$ is convex. We have that $\psi(t) = z^\top A^2 z + 2t(z^\top ABz) + t^2(z^\top B^2 z)$. Since the matrix B^2 is always positive semidefinite, it follows that $z^\top B^2 z \ge 0$, and hence indeed $\psi(t)$ is convex.

Example 2.3 The mapping $G(Z) := Z^{-1}$ is matrix convex on the set S_p^{++} of positive definite matrices. Indeed, for $A \in S_p^{++}$, $B \in S_p^{+}$ and $z \in R^p$, consider the real valued function $\psi(t) := z^\top G(A + tB)z$. We have that the second order derivative $\psi''(0)$ is equal to $2z^\top(A^{-1}BA^{-1}BA^{-1})z$. Since A^{-1} is positive definite and $(BA^{-1}z)^\top = z^\top A^{-1}B$, it follows that $\psi''(0) \ge 0$. Since this is true for an arbitrary positive definite matrix A, it follows that $\psi(t)$ is convex as long as $A + tB$ stays in S_p^{++}. This proves that Z^{-1} is matrix convex on the set S_p^{++}.

We need to notice that, in Examples 2.2 and 2.3, the variables are in the matrix space S_n, and actually S_n can be identified with $R^{\frac{n(n+1)}{2}}$. Hence, it can be consistent with Definition 2.1 and Proposition 2.2. For the above examples, we will develop the local nonsmooth optimization strategies suited for this new context, and show that they improve the situation considerably.

3 $\mathcal{VU}$-space decomposition

3.1 $\mathcal{VU}$-space

We study now the function of the largest eigenvalue, denoted by $f(x) = \lambda_1(A(x))$. In this subsection, we give the detailed structure of $\mathcal{VU}$-space for $\lambda_1(A(x))$.

Similar to the results in the literatures (Hiriart-Urruty and Ye 1995; Overton 1992), the subdifferential of the nonconvex maximum eigenvalue function $\partial(\lambda_1 \circ A)(x)$, by compositing the subdifferential of the convex component with the derivative of the smooth components, can be characterized in the following way.

Proposition 3.1 (Hiriart-Urruty and Ye 1995; Overton 1992)

$$\begin{aligned}\partial(\lambda_1 \circ A)(x) &= DA(x)^\star \partial\lambda_1(A(x)) \\ &= \left\{z \in R^m : z_k = \langle A_k, Q_1(x)Z(x)Q_1(x)^\top\rangle, k = 1, \ldots, m\right\},\end{aligned} \tag{3.1}$$

where $\partial\lambda_1(A(x)) = \{U(x) \in S_n : Z(x) \in S_r^+, \mathrm{tr}Z(x) = 1, U(x) = Q_1(x)Z(x)Q_1(x)^\top\}$ *and* $A_k(x) \equiv \frac{\partial A(x)}{\partial x_k}$ *for* $k = 1, \ldots, m$, $Q_1(x)$ *is an orthonormal basis of the first eigenspace for* $\lambda_1(A(x))$.

The relative interior of $\partial(\lambda_1 \circ A)(x)$ *has the expression*

$$\mathrm{ri}\partial(\lambda_1 \circ A)(x) = DA(x)^\star \mathrm{ri}\partial\lambda_1(A(x)), \tag{3.2}$$

where $\mathrm{ri}\partial\lambda_1(A(x)) = \{U(x) \in S_n : Z(x) \in S_r, Z(x) \succ 0, \mathrm{tr}Z(x) = 1, U(x) = Q_1(x)Z(x)Q_1(x)^\top\}$.

If $r = 1$, *then* $\lambda_1(A(x))$ *is differentiable at* x, *and*

$$\begin{aligned}\nabla(\lambda_1(A(x))) &= DA(x)^\star(Q_1(x)Q_1(x)^\top) \\ &= (\mathrm{tr}(Q_1(x)^\top A_1(x)Q_1(x)), \ldots, \mathrm{tr}(Q_1(x)^\top A_m(x)Q_1(x)))^\top.\end{aligned}$$

Proof Apply the chain rule given in Hiriart-Urruty and Lemaréchal (1993) to obtain (3.1) and the calculus rule to obtain (3.2). When $r = 1$, we can refer to Overton and Womersley (1995). □

Recall that by $[DA(x)]^{-1}$ we denote the multifunction with the graph inverse to that of $DA(x)$, whose form is

$$[DA(x)]^{-1}(Z) := \{h \in R^m : DA(x)h = Z\}.$$

In the sequel, we need the following assumption on the operator $DA(x)$.

Assumption 3.1 At a given point x, $DA(x)$ is surjective (onto), so the linear operator $DA(x)$ at x is invertible, i.e., $DA(x)^{-1}$ exists.

Note that the assumption is natural and important, it follows from Abraham et al. (1988), Bonnans and Shapiro (2000) that ensures the execution of space decomposition.

Remark 3.1 Notice that the mapping $DA(x)$ is a linear operator, which replies on some given point. However, in Ref. Oustry (1999) the operator $DA(x)$ reduces to the linear form $\mathcal{A}$, which is also the biggest difference. Next, we will see the results, such as $\mathcal{VU}$-space decomposition, first-order and second-order expansions, etc, not only depend on some given point x, but relate to the operator $DA(x)$ at this point.

When the mapping A reduces to a linear mapping: $A : R^m \ni x \mapsto A_0 + \mathcal{A} \cdot x = A_0 + \sum_{i=1}^m x_i A_i$, then $DA(x)$ reduces to the linear operator $\mathcal{A}$. Apparently, we find it doesn't depend on any point x, which means that its corresponding results are concerned with the operator $\mathcal{A}$, regardless of the selected point x. This is also the biggest difference with the nonlinear matrix convex mapping. So the linear mapping can be regarded as the special case of matrix convex form, this will lead to our conclusions being more general and complicated.

In this paper, we need to remember a major fact: a good model of λ_1 must consider the local behavior of all active constraints at x in the composition form. Geometry can be added: it suggests that fixing the multiplicity (i.e., the activity) of λ_1, the surface of activity is actually the smooth manifold $\mathcal{M}_r$. This point of view is the one adopted in Overton and Ye (1994). This gives us a geometrical interpretation of the subspaces $\mathcal{U}_f(x)$ and $\mathcal{V}_f(x)$.

Theorem 3.1 *(1) The subspaces $\mathcal{U}_f(x)$ and $\mathcal{V}_f(x)$ are respectively characterized by*

$$\begin{aligned}\mathcal{U}_f(x) &= DA(x)^{-1}\mathcal{U}_{\lambda_1}(A(x)) \\ &= DA(x)^{-1}\left\{U \in S_n : Q_1(x)^\top U Q_1(x) - \frac{1}{r}\mathrm{tr}(Q_1(x)^\top U Q_1(x)) I_r = 0\right\}\end{aligned} \tag{3.3}$$

and

$$\mathcal{V}_f(x) = DA(x)^\star \mathcal{V}_{\lambda_1}(A(x)) = DA(x)^\star\{Q_1(x)Z(x)Q_1(x)^\top : Z(x) \in S_r^+, \mathrm{tr}Z(x) = 0\}. \tag{3.4}$$

(2) If $r = 1$, then $\lambda_1(A(x))$ is a differentiable function, here we have $\mathcal{U}_f(x) = R^m$, and $\mathcal{V}_f(x) = \{0\}$.

(3) $\mathrm{int}\,\partial(\lambda_1 \circ A)(x) \neq \emptyset$, i.e., $\partial(\lambda_1 \circ A)(x)$ has full dimension, then we have $\mathcal{U}_f(x) = \{0\}$, and $\mathcal{V}_f(x) = R^m$.

Proof Taking the affine hull of the right- and left-hand sides in (3.1), because $DA(x)$ is a linear operator, by convex analysis (Rockafellar 1970), we obtain

$$\mathrm{aff}\partial(\lambda_1 \circ A)(x) = DA(x)^\star \mathrm{aff}\partial\lambda_1(A(x)).$$

By the definition of $\mathcal{VU}$-decomposition in (2.1), this gives the first part of (3.4). Write $\mathcal{U}_f(x) = \mathcal{V}_f(x)^\perp$ and deduce

$$\mathcal{U}_f(x) = \left\{u \in R^m : DA(x)(u) \in \mathcal{V}_{\lambda_1}(A(x))^\perp\right\},$$

which is the former one of (3.3), so next we only need to compute $\mathcal{V}_{\lambda_1}(A(x))$ and $\mathcal{U}_{\lambda_1}(A(x))$.

From the definition of the reference (Lemaréchal et al. 2000), we have $\mathcal{U}(x) = \{d \in R^n : f'(x; d) = -f'(x; -d)\}$, $\mathcal{V}(x) = \mathcal{U}(x)^{\perp}$, so here compute D which satisfies the equation

$$
\begin{aligned}
\mathcal{U}_{\lambda_1}(A(x)) &= \left\{D : \lambda_1'(A(x); D) = -\lambda_1'(A(x); -D)\right\} \\
&= \left\{D : \max_{G \in \partial\lambda_1(A(x))} \langle G, D\rangle = \min_{G \in \partial\lambda_1(A(x))} \langle G, D\rangle\right\} \\
&= \left\{D : \max_{\tilde{U}} \langle \tilde{U}, Q_1(x)^{\top} D Q_1(x)\rangle = \min_{\tilde{U}} \langle \tilde{U}, Q_1(x)^{\top} D Q_1(x)\rangle\right\} \\
&= \left\{D : Q_1(x)^{\top} D Q_1(x) = \tfrac{1}{r}\mathrm{tr}(Q_1(x)^{\top} D Q_1(x)) I_r\right\}.
\end{aligned}
$$

Therefore we obtain the remaining part of (3.3), likewise, make use of $\mathcal{V}_{\lambda_1}(A(x)) = \mathcal{U}_{\lambda_1}(A(x))^{\perp}$, (3.4) also holds. □

3.2 The $\mathcal{U}$-Lagrangian function of $\lambda_1(A(x))$

Take $g^* \in \mathrm{ri}\partial(\lambda_1 \circ A)(x^*)$ and define the $\mathcal{U}$-Lagrangian of f at (x^*, g^*) according to (2.1); in the following we denote it by $L_{\mathcal{U},f}(x^*, g^*; \cdot)$. From item 2 of Theorem 2.1, $L_{\mathcal{U},f}(x^*, g^*; \cdot)$ is differentiable at $u = 0$. We can prove the following composition rule.

Theorem 3.2 *Let $G^* \in \mathrm{ri}\partial\lambda_1(A(x^*))$ be such that $g^* = DA(x^*)^{\star} \cdot G^*$. Then,*

$$
\nabla L_{\mathcal{U},f}(x^*, g^*; 0) = [\mathrm{proj}_{\mathcal{U}_f(x^*)} \circ DA(x^*)^{\star} \circ \mathrm{proj}^*_{\mathcal{U}_{\lambda_1}(A(x^*))}] \cdot \nabla L_{\mathcal{U},f}(A(x^*), G^*; 0), \tag{3.5}
$$

where $\mathcal{U}_f(x)$ is given by (3.3).

Proof Due to (3.2), we can always find some G^*, satisfying the assumption. Using (2.5),

$$
\begin{aligned}
\nabla L_{\mathcal{U},f}(x^*, g^*; 0) &= \mathrm{proj}_{\mathcal{U}_f(x^*)} g^* \\
&= \mathrm{proj}_{\mathcal{U}_f(x^*)} (DA(x^*)^{\star} \cdot G^*) \\
&= \mathrm{proj}_{\mathcal{U}_f(x^*)} \circ DA(x^*)^{\star} \cdot (\mathrm{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))} G^* \oplus \mathrm{proj}_{\mathcal{V}_{\lambda_1}(A(x^*))} G^*) \\
&= \mathrm{proj}_{\mathcal{U}_f(x^*)} \circ DA(x^*)^{\star} \cdot (\mathrm{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))} G^* \oplus 0),
\end{aligned}
$$

where the third equal sign holds above, because the operator $DA(x^*)$ is linear.

Since according to (3.4), we have

$$
DA(x^*)^{\star} \cdot \left(0 \oplus \mathrm{proj}_{\mathcal{V}_{\lambda_1}(A(x^*))} G^*\right) \in \mathcal{V}_f(x^*) = \mathcal{U}_f(x^*)^{\perp},
$$

the last equality in the above formula holds.

At last, from $\nabla L_{\mathcal{U},f}(A(x^*), G^*; 0) = \text{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))} G^*$ and the definition of the adjoint operator,

$$\text{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))} G^* \oplus 0 = \text{proj}^*_{\mathcal{U}_{\lambda_1}(A(x^*))} \nabla L_{\mathcal{U},f}(A(x^*), G^*, 0),$$

and the conclusion (3.5) holds. □

We next introduce the subspace

$$\mathcal{H} := \{Z \in S_r : \text{tr} Z = 0\}, \tag{3.6}$$

and consider the map

$$\Phi : B(A(x^*), \delta) \ni A \mapsto Q_1(A(x))^\top A(x) Q_1(A(x)) - \frac{1}{r}\text{tr}\,(Q_1(A(x))^\top A(x) Q_1(A(x))) I_r \in \mathcal{H}. \tag{3.7}$$

Here, we would like to identify a characteristic C^∞-manifold. A natural idea is to test the set of vectors $x \in R^m$ such that $\lambda_1(A(x))$ has a fixed multiplicity r; namely to consider $A^{-1}(\mathcal{M}_r)$. The difficulty is that, even in the affine case, $A^{-1}(\mathcal{M}_r)$ may be nonsmooth. To ensure that $A^{-1}(\mathcal{M}_r)$ is a smooth manifold in a neighborhood of x^*, we need to assume that A^{-1} is transversal to $\mathcal{M}_r$, i.e.,

Definition 3.1 We say that the transversality condition **(T)** holds at $x^* \in A^{-1}(\mathcal{M}_r)$ if

$$\textbf{(T)} \qquad range DA(x^*) + \mathcal{U}_{\lambda_1}(A(x^*)) = S_n. \tag{3.8}$$

Remark 3.2 Transversality assumption (T) used by Shapiro and Fan (cf, Shapiro and Fan 1995; Shapiro 1997), in fact, is equivalent to the linear independence condition, which is some kind of constraint qualification. In addition, the condition (T) is called the constraint *nondegenerated* condition in semidefinite programming (Bonnans and Shapiro 2000). It is an analogue of the condition of linear independence of the gradients of active constraints used in nonlinear programming. It is a well known phenomenon that for very large linear programming quite often an optimal solution tends to happen at numerically degenerate points. Transversality is also stable under small perturbations and a generic property. It has been extensively used for sensitivity and stability analysis in optimization and variational inequalities.

Besides the guarantee of a smooth manifold $\mathcal{W}_r := A^{-1}(\mathcal{M}_r)$ around a point x, transversality assures that $\lambda_1(A(x))$ is partly smooth at x relative to $\mathcal{W}_r$ defined by Lewis (2003), meanwhile, the quadratic convergence is guaranteed in a later algorithm.

Next we obtain a local equation of $\mathcal{W}_r$ via a simple composition rule.

Theorem 3.3 *If (T) is satisfied at x^*, then there exists $\rho > 0$ such that $\varphi(x) = 0$, where $\varphi : B(x^*, \rho) \ni x \mapsto \Phi(A(x)) \in S_r$ and Φ given by (3.7), is a local equation of $\mathcal{W}_r \cap B(x^*, \rho)$. Moreover, for all $x \in B(x^*, \rho)$, we have*

$$\text{T}_{\mathcal{W}_r}(x) = \ker D\varphi(x).$$

In addition, if $g^* \in \mathrm{ri}\partial f(x^*)$*, then*

(1) the subspaces $\mathcal{U}_f(x^*)$ *and* $\mathcal{V}_f(x^*)$ *are respectively the tangent and normal spaces to* $\mathcal{W}_r$ *at* x^**, i.e.,*

$$\mathcal{U}_f(x^*) = T_{\mathcal{W}_r}(x^*) = \ker(D\varphi(x^*));$$
$$\mathcal{V}_f(x^*) = N_{\mathcal{W}_r}(x^*) = \mathrm{range}(D\varphi(x^*)^\star).$$

(2) there exists $\rho > 0$ *and a* C^∞*-mapping* $v : \mathcal{U}_f(x^*) \cap B(0,\rho) \to \mathcal{V}_f(x^*)$ *such that the mapping*

$$p_{x^*} : \mathcal{U}_f(x^*) \cap B(0,\rho) \ni u \mapsto x^* + u \oplus v(u), \tag{3.9}$$

is a C^∞ *tangential parametrization of the submanifold* $\mathcal{W}_r$*.*

Proof Because the mapping of (3.7) defines a local equation of $\mathcal{M}_r$ and the transversality condition (T) is satisfied, there exists $\rho > 0$ such that $\Phi \circ A$ defines a local equation of $\mathcal{W}_r \cap B(x^*,\rho)$.

From Ref. Abraham et al. (1988), it holds that, for all $x \in B(x^*,\rho)$, $T_{\mathcal{W}_r}(x) = \ker D\varphi(x)$.

(1) Because (T) holds at x^*, we have $T_{\mathcal{W}_r}(x) = [DA(x)]^{-1}T_{\mathcal{M}_r}(A(x))$. By the formula of (3.3), we get $\mathcal{U}_f(x^*) = T_{\mathcal{W}_r}(x^*)$ and $\mathcal{V}_f(x^*) = N_{\mathcal{W}_r}(x^*)$.
(2) We can directly apply Theorem 2.2 (1) to obtain the C^∞ mapping $v : \mathcal{U}_f(x^*) \cap B(0,\rho) \to \mathcal{V}_f(x^*)$ and (2) to get the tangential parametrization of the submanifold p_{x^*}. □

The mapping (3.9) is a geometrical interpretation that $\mathcal{U}_f(x)$ is tangent at x^* to the ridge, in our context this geometrical set (3.9) coincides in a neighborhood of x^* with $\mathcal{W}_r$ when $g^* \in \mathrm{ri}\partial f(x^*)$. The next result shows a nice interpretation of $w(\cdot)$ in (2.3), which makes a local description of the surface $x^* + u \oplus w(u)$.

Theorem 3.4 *Suppose that the condition **(T)** is satisfied at* x^* *and take* $g^* \in \mathrm{ri}\,\partial f(x^*)$*. Then there exists* $\rho > 0$ *such that for all* $u \in B(0,\rho) \subset \mathcal{U}_f(x^*)$*, the set* $w(u)$ *is a singleton:*

$$w(u) = \{v(u)\} \text{ for all } u \in B(0,\rho),$$

where $v(\cdot)$ *is the* C^∞*-mapping defined in Theorem 3.3 (2).*

Proof Set $u \in \mathcal{U}(x^*)$, $v \in w(u)$, and

$$G \in [DA(x)^\star]^{-1}\left[\partial\lambda_1 \circ A\left(x^* + u \oplus v\right) \cap \left(g^* + \mathcal{U}(x^*)\right)\right] \cap \partial\lambda_1\left(A(x^* + u \oplus v)\right).$$

From the formulation of $\partial\lambda_1 \circ A(x^* + u \oplus v)$, we have $G = Q_1(x^* + u \oplus v)ZQ_1(x^* + u \oplus v)^\top$, $\mathrm{tr}Z = 1$, which implies the following complementarity condition holds

$$\left(\lambda_1\left(A\left(x^* + u \oplus v\right)\right)I_n - A\left(x^* + u \oplus v\right)\right)G = 0.$$

Through the knowledge of matrix analysis, we get the rank condition

$$\operatorname{rank}\left(\lambda_1\left(A\left(x^*+u\oplus v\right)\right)I_n - A\left(x^*+u\oplus v\right)\right) + \operatorname{rank}G \le n. \tag{3.10}$$

Furthermore, at $u=0$, $G=G^*\in \operatorname{ri}\partial\lambda_1(A(x^*+u\oplus v))$, and we find the following strict complementarity condition holds by (3.2):

$$\operatorname{rank}\left(\lambda_1\left(A\left(x^*\right)\right)I_n - A\left(x^*\right)\right) = n-r \text{ and } \operatorname{rank}G^* = r.$$

Because of the continuity of eigenvalues with (2.5) and (2.6), there is a positive number $\rho>0$ that satisfies

$$\begin{aligned}&\operatorname{rank}\left(\lambda_1\left(A\left(x^*+u\oplus v\right)\right)I_n - A\left(x^*+u\oplus v\right)\right) \ge n-r \;\text{ and }\; \operatorname{rank}G \ge r\\ &\quad\text{for all } u\in B(0,\rho) \text{ and}\\ &\quad\text{all } (v,G)\in w(u)\left[DA(x)^\star\right]^{-1}\left[\partial\lambda_1\circ A\left(x^*+u\oplus w(u)\right)\cap g^*+\mathcal{U}(x^*)\right].\end{aligned}$$

Combining the above formula with the inequality (3.10), we obtain

$$\begin{aligned}&\operatorname{rank}\left(\lambda_1\left(A\left(x^*+u\oplus v\right)\right)I_n - A\left(x^*+u\oplus v\right)\right) = n-r \;\text{ and }\; \operatorname{rank}G = r\\ &\quad\text{for all } u\in B(0,\eta) \text{ and}\\ &\quad\text{all } (v,G)\in w(u)\left[DA(x)^\star\right]^{-1}\left[\partial\lambda_1\circ A\left(x^*+u\oplus w(u)\right)\cap g^*+\mathcal{U}(x^*)\right],\end{aligned}$$

which is the strict complementarity condition needed. Then, take ρ small enough, we have

$$x^*+u\oplus v\subset B(x^*,\delta)\cap\mathcal{W}_r,$$

so we can apply Theorem 2.2 (2) and to derive the formula $w(u)=\{v(u)\}$, the proof is finished. $\square$

Now we follow the path $p_{x^*}(u)\in\mathcal{W}_r$. On the manifold $\mathcal{W}_r$ and close enough to x^*, the subspace $E_{tot}(\cdot)$ spanned by the first r eigenvectors and $E_1(\cdot)$ coincide. Hence one chooses an orthonormal basis mapping in a neighborhood of $u=0$:

$$\mathcal{U}(x^*)\ni u\to Q_1(p_{x^*}(u)) := Q_{tot}(p_{x^*}(u)),$$

where the columns of $Q_{tot}(\cdot)$ form an orthonormal basis of $E_{tot}(\cdot)$.

The above theorem shows that the $\mathcal{U}$-Lagrangian is useful for the analytic construction of the implicitly-defined v. Next we show the existence of $\nabla^2 L_{\mathcal{U},f}(x^*,g^*;0)$, the so-called $\mathcal{U}$-Hessian matrix of λ_1 at x^*. So this leads us to our main result.

Theorem 3.5 *Assume that the transversality condition* ***(T)*** *is satisfied at x^* and take $g^*\in\operatorname{ri}\partial f(x^*)$. Then the $\mathcal{U}$-Lagrangian function $L_{\mathcal{U},f}(x^*,g^*;\cdot)$ of f is C^∞ in a neighborhood of $u=0$. In particular, at $u=0$,*

$$\nabla^2 L_{\mathcal{U},f}(x^*, g^*; 0) = \mathrm{proj}_{\mathcal{U}_f(x^*)} \circ DA(x^*)^\star \circ H(A(x^*), G^*) \circ DA(x^*) \circ \mathrm{proj}^*_{\mathcal{U}_f(x^*)}, \quad (3.11)$$

where G^* is the unique subgradient of $\partial\lambda_1(A(x^*))$ such that $g^* = DA(x^*)^\star G^*$ and the operator $H(A(x^*), G^*)$ is the symmetric positive semidefinite operator defined by

$$H(A(x^*), G^*) \cdot Y = G^* Y[\lambda_1^* I_n - A(x^*)]^\dagger + [\lambda_1^* I_n - A(x^*)]^\dagger Y G^*. \quad (3.12)$$

This can also be written

$$\nabla^2 L_{\mathcal{U},f}(x^*, g^*; 0) = B(x^*)^* \circ \nabla^2 L_{\mathcal{U}}(A(x^*), G^*; 0) \circ B(x^*), \quad (3.13)$$

where $B(x^*) = \mathrm{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))} \circ DA(x^*) \circ \mathrm{proj}^*_{\mathcal{U}_f(x^*)}$ and $\nabla^2 L_{\mathcal{U}}(A(x^*), G^*; 0) = \mathrm{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))} \circ H(A(x^*), G^*) \circ \mathrm{proj}^*_{\mathcal{U}_{\lambda_1}(A(x^*))}$, and $\mathcal{U}_f(x^*)$ is given by (3.3).

Proof Because $\lambda_1(A(x)) = \hat{\lambda}(A(x)) := \frac{1}{r}\sum_{i=1}^r \lambda_i(A(x))$ for all $x \in \mathcal{W}_r$ close enough to x^*, Theorem 3.4 gives

$$L_{\mathcal{U},f}(x^*, g^*; u) = \lambda_1(p_{x^*}(u)) - \langle g^*, v(u)\rangle_{\mathcal{V}_f(x^*)} \text{ for all } u \in B(0,\rho). \quad (3.14)$$

Then $L_{\mathcal{U}}(x^*, g^*; \cdot)$ is C^∞ on $B(0,\rho)$. Next we proceed in two steps.

(1) Make use of (2.4) and (3.1) to obtain

$$\partial L_{\mathcal{U}}(u) = \mathrm{proj}_{\mathcal{U}_f(x^*)} DA(x)^\star \Big\{ Q_1\left(p_{x^*}(u)\right) Z Q_1\left(p_{x^*}(u)\right)^\top : Q_1\left(p_{x^*}(u)\right) Z Q_1\left(p_{x^*}(u)\right)^\top - G^* \in \mathcal{U}(A^*),\, Z \in S_r^+,\, \mathrm{tr}Z = 1 \Big\}.$$

From Theorem 3.3, we have $\mathcal{U}_f(x^*) = T_{\mathcal{W}_r}(x^*)$, i.e., $\mathcal{U}_f(x^*) = \ker D\varphi(x^*)$. Then by Theorem 3.1 we get

$$D\varphi(x^*) \cdot DA(x)^{-1}\left[\left(Q_1(p_{x^*}(u)) Z Q_1(p_{x^*}(u))^\top - G^*\right)\right] = 0 \text{ and } \mathrm{tr}Z = 1. \quad (3.15)$$

Now according to the condition that Z satisfies, set $Z = \frac{1}{r}I_r + \Xi$, where Ξ is an element of $\mathcal{H}$ defined in (3.6). Then introducing

$$D\varphi\left(p_{x^*}(u)\right)^* : \mathcal{H} \ni \Xi \mapsto DA(x)^{-1}\left(Q_1\left(p_{x^*}(u)\right) \Xi Q_1\left(p_{x^*}(u)\right)^\top\right) \in R^m,$$

we combine the above formula and (3.15)

$$\begin{aligned} D\varphi(x^*) \circ D\varphi(p_{x^*}(u))^* \cdot \Xi &= D\varphi(x^*) \circ DA(x)^{-1}\left(Q_1\left(p_{x^*}(u)\right)\Xi Q_1\left(p_{x^*}(u)\right)^\top\right) \\ &= D\varphi(x^*) \cdot \left[DA(x)^{-1}G^*\right] \\ &\quad - \frac{1}{r}D\varphi(x^*)\left[DA(x)^{-1}\left(Q_1\left(p_{x^*}(u)\right)Q_1\left(p_{x^*}(u)\right)^\top\right)\right]. \end{aligned} \tag{3.16}$$

Because the mapping $D\varphi(x^*)$ is onto, $D\varphi(x^*) \circ D\varphi(x^*)^*$ is invertible. $D\varphi(x^*) \circ D\varphi(p_{x^*}(u))^*$ is also invertible for u small enough, making use of the continuity of $D\varphi(x^*) \circ D\varphi(x^*)^*$. Then we can get the unique solution $\Xi(u)$ through inverting the two sides of (3.16), and define the C^∞-mapping

$$\mathcal{U}_f(x^*) \cap B(0,\rho) \ni u \mapsto Z(u) := \frac{1}{r}I_r + \Xi(u), \tag{3.17}$$

so

$$\nabla L_{\mathcal{U},f}\left(x^*, g^*; u\right) = \mathrm{proj}_{\mathcal{U}_f(x^*)} \circ DA(x)^\star \cdot \left(Q_1\left(A\left(p_{x^*}(u)\right)\right)Z(u)Q_1\left(A\left(p_{x^*}(u)\right)\right)^\top\right). \tag{3.18}$$

(2) We differentiate the formula (3.18) at $u = 0$ to get the second-order term; since we apply a fixed linear operator $\mathrm{proj}_{\mathcal{U}_f(x^*)}$, we obtain the sum of the three terms. One of these terms is

$$\mathrm{proj}_{\mathcal{U}_f(x^*)} \circ DA(x^*)^\star \cdot \left(Q_1(A(x^*))\left[DZ(0)\cdot d\right]Q_1(A(x^*))^\top\right), \textit{for } d \in \mathcal{U}_{f_k}(x^*)$$

which is zero: indeed $DZ(0) \cdot d = D\Xi(0) \cdot d \in \mathcal{H}$ from (3.17) and therefore $DA(x^*)^\star \cdot (Q_1(A(x^*))[DZ(0)\cdot d]Q_1(A(x^*))^\top) \in \mathcal{V}_f(x^*)$. Also, we have

$$Dp_{x^*}(0) \cdot d = d + DV(0) \cdot d = d.$$

Using (3.8) of Shapiro and Fan (1995), we obtain the required result.

Finally, because of (3.3) we have

$$\mathrm{range}\left(DA(x^*) \circ \mathrm{proj}^*_{\mathcal{U}_f(x^*)}\right) = \mathrm{proj}^*_{\mathcal{U}_{\lambda_1}(A(x^*))}\mathcal{U}_{\lambda_1}(A(x^*)).$$

In virtue of $\nabla^2 L_{\mathcal{U}}(A(x^*), G^*; 0) = \mathrm{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))} \circ H(A(x^*), G^*) \circ \mathrm{proj}^*_{\mathcal{U}_{\lambda_1}(A(x^*))}$ the operator $H(A(x^*), G^*)$ can be replaced in (3.11) by

$$\mathrm{proj}^*_{\mathcal{U}_{\lambda_1}(A(x^*))} \circ \nabla^2 L_{\mathcal{U},f}\left(A(x^*), G^*; 0\right) \circ \mathrm{proj}_{\mathcal{U}_{\lambda_1}(A(x^*))},$$

and we finish the proof. □

The operator induced by $H(A^*, G^*)$ in the subspace $\mathcal{U}_{\lambda_1}(A(x^*))$ is called the $\mathcal{U}$-*Hessian* of λ_1 at pair (A^*, G^*). It collects the relevant second-order information on λ_1 along the manifold $\mathcal{M}_r$. A second-order-like expansion of f on $\mathcal{W}_r$ is obtained from the $\mathcal{U}$-Hessian matrix.

Corollary 3.1 *Under the assumptions of Theorem 3.5, we have for all $d \in R^m$, such that $x^* + d \in \mathcal{W}_r$ and $d \to 0$,*

$$\lambda_1 \circ A(x^* + d) = \lambda_1 \circ A(x^*) + \langle g^*, d\rangle + \frac{1}{2}\left\langle \mathrm{proj}_{\mathcal{U}_f(x^*)} d, \nabla^2 L_{\mathcal{U},f}\left(x^*, g^*; 0\right) \cdot \left(\mathrm{proj}_{\mathcal{U}_f(x^*)} d\right)\right\rangle_{\mathcal{U}_f(x^*)} + o(\|d\|^2). \tag{3.19}$$

Proof Let d be small enough such that $x^* + d \in \mathcal{W}_r$, employ (2.11) and set $u = \mathrm{proj}_{\mathcal{U}_f(x^*)} d$, $v = v(u) = \mathrm{proj}_{\mathcal{V}_f(x^*)} d$, apply Theorem 3.5:

$$\begin{aligned} L_{\mathcal{U}}(x^*, g^*; u) &= \lambda_1 \circ A(x^*) + \nabla L_{\mathcal{U}}(x^*, g^*; 0) \cdot u \\ &\quad + \tfrac{1}{2} u \cdot \nabla^2 L_{\mathcal{U}}\left(x^*, g^*; 0\right) \cdot u + o\left(\|u\|^2\right) \\ &= \lambda_1 \circ A\left(x^* + u \oplus v(u)\right) - \mathrm{proj}_{\mathcal{V}_f(x^*)} g^* \cdot v(u). \end{aligned}$$

Hence,

$$\begin{aligned} \lambda_1 \circ A(x^* + u \oplus v(u)) &= \lambda_1 \circ A(x^*) + \nabla L_{\mathcal{U}}(x^*, g^*; 0) \cdot u + \mathrm{proj}_{\mathcal{V}_f(x^*)} g^* \cdot v(u) \\ &\quad + \tfrac{1}{2} u \cdot \nabla^2 L_{\mathcal{U}}(x^*, g^*; 0) \cdot u + o(\|u\|^2) \\ &= \lambda_1 \circ A(x^*) + \langle g^*, u \oplus v(u)\rangle \\ &\quad + \tfrac{1}{2} u \cdot \nabla^2 L_{\mathcal{U}}(x^*, g^*; 0) \cdot u + o(\|u\|^2). \end{aligned}$$

Finally, recall (2.10), we have $v = o(\|u\|^2) = o(\|d\|^2)$, and we are done. □

4 Smooth trajectories and algorithm

4.1 Smooth trajectories and second-order expansion

There is also another purpose of introducing the transversality condition (**T**), that is to identify the trajectories along which $\lambda_1(\cdot)$ behaves in a smooth manner. For the maximum eigenvalue function, this condition (**T**) is assumed by many authors to obtain second-order developments (see Overton and Womersley 1995; Shapiro and Fan 1995). Next we will show how to find a smooth trajectory that is tangent to $\mathcal{U}$ using the structure of λ_1 and an implicit function theorem to parameterize the trajectory based on u.

First we give some signs appearing in this section. Let $q_1(x), \ldots, q_n(x)$ be a set of orthonormal eigenvectors of $A(x)$ corresponding to the eigenvalues $\lambda_1(x), \ldots, \lambda_n(x)$. With the mapping $A : R^m \to S_n$ are the associated symmetric matrices, $A_l(x) = \frac{\partial A(x)}{\partial x_l}$, $l = 1, \ldots, m$, and m-dimensional vectors

$$v_{ij}(x) = \left(q_i(x)^\top A_1(x) q_j(x), \cdots, q_i(x)^\top A_m(x) q_j(x)\right)^\top, \quad i, j = 1, \ldots, n.$$

Set

$$I_1 := \{1, 2, \ldots, r\} \quad and \quad I_2 := \{(k, l) \in I_1 \times I_1 : k < l\},$$

so $|I_1| = r$, $|I_2| = \frac{r(r-1)}{2}$. According to the transversality conditions, we have the following lemma for the mapping A with respect to the smooth manifold $\mathcal{W}_r$.

Lemma 4.1 (Shapiro and Fan 1995) *Suppose that $A(x) \in \mathcal{W}_r$. Then the transversality condition holds if and only if the vectors $v_{kl}(x), v_{ii}(x) - v_{rr}(x)$ are linearly independent,* $1 \le k < l \le r$, $i = 1, \ldots, r-1$.

In view of the definition of the eigenvalues and eigenvectors, it is easy to find that

$$\begin{aligned} \lambda_1(A(x)) &= q_i^\top(x) A(x) q_i(x),\ i \in I_1, \\ 0 &= q_k^\top(x) A(x) q_l(x),\ (k, l) \in I_2 \end{aligned} \tag{4.1}$$

and

$$\begin{aligned} q_i^\top(x) q_i(x) &= 1,\ i \in I_1, \\ q_k^\top(x) q_l(x) &= 0,\ (k, l) \in I_2. \end{aligned} \tag{4.2}$$

According to the above lemma, we have the following result:

Lemma 4.2 *The partial derivative of the formula* (4.1) *with respect to the variable x_j, $j = 1, \ldots, m$ can be written as*

$$\begin{aligned} \frac{\partial \lambda_1(A(x))}{\partial x_j} &= q_i^\top(x) \frac{\partial A(x)}{\partial x_j} q_i(x),\ i \in I_1, \\ 0 &= q_k^\top(x) \frac{\partial A(x)}{\partial x_j} q_l(x),\ (k, l) \in I_2. \end{aligned} \tag{4.3}$$

Proof For $j = 1, \ldots, m$, differentiate the first formula of (4.1) with respect to x_j, we obtain

$$\begin{aligned} \frac{\partial \lambda_1(A(x))}{\partial x_j} &= q_i^\top(x) \frac{\partial A(x)}{\partial x_j} q_i(x) + \left\{ \frac{\partial q_i(x)}{\partial x_j}^\top A(x) q_i(x) + q_i^\top(x) A(x) \frac{\partial q_i(x)}{\partial x_j} \right\} \\ &= q_i^\top(x) \frac{\partial A(x)}{\partial x_j} q_i(x) + \left\{ \lambda_1(A(x)) \frac{\partial q_i(x)}{\partial x_j}^\top q_i(x) + \lambda_1(A(x)) q_i^\top(x) \frac{\partial q_i(x)}{\partial x_j} \right\}, i \in I_1, \end{aligned}$$

where the second equality of the above formula holds, because of $A(x) q_i(x) = \lambda_1(A(x)) q_i(x)$.

Differentiating the first formula of (4.2), having the constant right side, gives

$$\frac{\partial q_i(x)}{\partial x_j}^\top q_i(x) + q_i^\top(x) \frac{\partial q_i(x)}{\partial x_j} = 0,$$

hence, the desired result is obtained. Similarly, we can get the second formula of (4.3). $\square$

For the maximum eigenvalue function λ_1, a second-order implicit function theorem gives a smooth primal track that is tangent to $\mathcal{U}$. In addition, we give the notation: $\bar{U}$ and $\bar{V}$ are respectively denoted the basis matrices for the subspaces $\mathcal{U}$ and $\mathcal{V}$, $u \oplus v = \bar{U}u + \bar{V}v$.

Theorem 4.1 *If the transversality condition **(T)** holds at $\bar{x}$, then for all u small enough,*

(i) the nonlinear equation system, with variables u and v

$$\begin{cases} q_i^\top(\bar{x} + \bar{U}u + \bar{V}v)A(\bar{x} + \bar{U}u + \bar{V}v)q_i(\bar{x} + \bar{U}u + \bar{V}v) \\ \quad -q_1^\top(\bar{x} + \bar{U}u + \bar{V}v)A(\bar{x} + \bar{U}u + \bar{V}v)q_1(\bar{x} + \bar{U}u + \bar{V}v) = 0, \ i \in I_1, i \neq 1, \\ q_k^\top(\bar{x} + \bar{U}u + \bar{V}v)A(\bar{x} + \bar{U}u + \bar{V}v)q_l(\bar{x} + \bar{U}u + \bar{V}v) = 0, \qquad (k,l) \in I_2 \end{cases} \tag{4.4}$$

has a unique solution $v = v(u)$, where $v : R^{dim\mathcal{U}} \to R^{dim\mathcal{V}}$ is a C^1-function;

(ii) the trajectory $\mathcal{X}(u) := \bar{x} + \bar{U}u + \bar{V}v(u)$ has a continuous Jacobian

$$J\mathcal{X}(u) := \bar{U} - \bar{V}(V(u)^\top \bar{V})^{-1}V(u)^\top \bar{U},$$

where

$$\begin{aligned} V(u) := &\Big[\nabla\left(q_i^\top(\mathcal{X}(u))\,A(\mathcal{X}(u))\,q_i(\mathcal{X}(u)) - q_1^\top(\mathcal{X}(u))\,A(\mathcal{X}(u))\,q_1(\mathcal{X}(u))\right)_i \\ &\qquad \cup \nabla\left(q_k^\top(\mathcal{X}(u))\,A(\mathcal{X}(u))\,q_l(\mathcal{X}(u))\right)_{(k,l)}\Big]; \\ = &\Big[\left\{v_{ii}(\mathcal{X}(u)) - v_{11}(\mathcal{X}(u))\right\}_{i\neq 1, i \in I_1} \cup \left\{v_{kl}(\mathcal{X}(u))\right\}_{(k,l)\in I_2}\Big] \end{aligned}$$

(iii) in particular, $v(0) = 0$, $\mathcal{X}(0) = \bar{x}$, $V(0) = \bar{V}$, $Jv(0) = 0$, and $J\mathcal{X}(0) = \bar{U}$;
(iv) $v(u) = O(|u|^2)$ and the trajectory $\mathcal{X}(u)$ is tangent to $\mathcal{U}$ at $\mathcal{X}(0) = \bar{x}$;
(v) $f(\mathcal{X}(u)) = q_i(\mathcal{X}(u))$ for $i = 1, \ldots, r$.

Proof (i) According to Lemma 4.2, differentiating the left hand of (4.4) with respect to v gives

$$\begin{cases} [v_{ii}(\bar{x} + \bar{U}u + \bar{V}v) - v_{11}(\bar{x} + \bar{U}u + \bar{V}v)]^\top \bar{V}, \ i \in I_1, i \neq 1, \\ v_{kl}^\top(\bar{x} + \bar{U}u + \bar{V}v)\bar{V}, \qquad (k,l) \in I_2. \end{cases}$$

This Jacobian at $(u, v) = (0, 0)$ is $\bar{V}^\top \bar{V}$. Because of the transversality condition and Lemma 4.1, it is nonsingular. With regard to u, there is also a Jacobian, hence, by the implicit function theorem, there exists a C^1 function $v(u)$ defined on a neighborhood of $u = 0$ satisfying $v(0) = 0$.

(ii) By (i), we find $v(u)$ is C^1, so the Jacobians $Jv(u)$ and $J\mathcal{X}(u)$ exist and are continuous. Differentiating the following equation system with respect to u

$$\begin{cases} q_i^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_i(\mathcal{X}(u)) - q_1^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_1(\mathcal{X}(u)) = 0, \ i \in I_1, i \neq 1, \\ q_k^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_l(\mathcal{X}(u)) = 0, \qquad (k,l) \in I_2, \end{cases}$$

we obtain that

$$\begin{cases} [v_{ii}(\mathcal{X}(u)) - v_{11}(\mathcal{X}(u))]^\top J(\mathcal{X}(u)) = 0, & i \in I_1, i \neq 1, \\ v_{kl}^\top(\mathcal{X}(u)) J(\mathcal{X}(u)) = 0, & (k,l) \in I_2, \end{cases}$$

or in matrix form, $V(u)^\top J(\mathcal{X}(u)) = 0$. Using the expression $J(\mathcal{X}(u)) = \bar{U} + \bar{V} J v(u)$, we have that

$$V(u)^\top(\bar{U} + \bar{V} J v(u)) = 0.$$

In the light of the continuity of $V(u)$, $V(u)^\top \bar{V}$ is nonsingular, i.e., $V(u)^\top \bar{V}$ is invertible. Hence

$$J v(u) = -(V(u)^\top \bar{V})^{-1} V(u)^\top \bar{U}.$$

Furthermore, $V(u)$ is C^1 because $q_i^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_i(\mathcal{X}(u)), q_k^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_l(\mathcal{X}(u))$ are C^1, then $Jv(u)$ is C^1. Thus $\mathcal{X}(u)$ and $v(u)$ are C^2.

(iii) By the definition of $\mathcal{VU}$ space decomposition, we have $\mathcal{V} \perp \mathcal{U}$. Hence $\bar{V}^\top \bar{U} = 0$. So $Jv(0) = 0$ and $J\mathcal{X}(0) = \bar{U}$.

(iv) By the Taylor expansion, since $Jv(0) = 0$, we have

$$v(u) = v(0) + Jv(0)u + o(\|u\|) = 0 + 0 \cdot u + o(\|u\|) = o(\|u\|).$$

So the trajectory

$$\mathcal{X}(u) = \mathcal{X}(0) + J\mathcal{X}(0)u + o(\|u\|) = \bar{x} + \bar{U}u + o(\|u\|),$$

i.e., $\mathcal{X}(u)$ is tangent to $\mathcal{U}$ at $\mathcal{X}(0) = \bar{x}$.

(v) It follows from the definition of $\mathcal{X}(u)$. □

Next let us see an example to illustrate the existence of the smooth trajectory.

Example For an function $f(x) = \lambda_1(\mathcal{F}(x))$ on R^2, where

$$\mathcal{F}(x) := \begin{pmatrix} x_1^2 & x_2 \\ x_2 & 0 \end{pmatrix},$$

the $\mathcal{VU}$-space decomposition at the point $\bar{x} = (0, 0)$ is presented as

$$\mathcal{V} = \text{lin}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \quad \textit{and} \quad \mathcal{U} = \text{lin}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right),$$

so we can take $\bar{U} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\bar{V} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Through solving (4.4) with $\bar{U}$ and $\bar{V}$,

$$\begin{cases} f_1(\bar{x} + \bar{U}u + \bar{V}v) - f_0(\bar{x} + \bar{U}u + \bar{V}v) = 0 \\ \varphi_1(\bar{x} + \bar{U}u + \bar{V}v) = 0, \end{cases}$$

where the relevant C^2 functions are $f_0(x) = x_1^2$, $f_1(x) = 0$, $\varphi_1(x) = x_2$. We can easily work out the solution $v(u) = 0$, so the smooth trajectory $x(u) = \bar{x} + \bar{U}u + \bar{V}v = (u, 0)^\top$.

The next result shows the associated dual object $\alpha = \alpha(u)$ that is also a smooth function of the variable u. It is useful to express the gradient and Hessian of $L_{\mathcal{U}}$ as combinations of the gradients and Hessians of the primal function.

Proposition 4.1 *Suppose transversality condition (T) holds with the trajectory $\mathcal{X}(u) = \bar{x} + \bar{U}u + \bar{V}v(u)$ and relative to $\bar{g} \in \partial\lambda_1 \circ A(\bar{x})$, for each u small enough, the linear equation system with the variables $\{\alpha_i\}$, $\{\alpha_j\}$*

$$\begin{cases} \bar{V}^\top \left[\sum_{i\in I_1} \alpha_i v_{ii}(\mathcal{X}(u)) + \sum_{(k,l)\in I_2} \alpha_j v_{kl}(\mathcal{X}(u)) \right] = \bar{V}^\top \bar{g} \in R^{m-1}, \\ \sum_{i\in I_1} \alpha_i = 1, \end{cases}$$

has a unique solution $\alpha = \alpha(u)$, expressed as

$$\begin{aligned} \{\alpha_i(u)\}_{i\neq 1, i\in I_1} \cup \{\alpha_j(u)\} &= (\bar{V}^\top V(u))^{-1} \bar{V}^\top (\bar{g} - v_{11}(\mathcal{X}(u)), \\ \alpha_1(u) &= 1 - \sum_{i\neq 1, i\in I_1} \alpha_i(u). \end{aligned}$$

In particular, $\alpha_i(0) = \bar{\alpha}_i$.

Proof For the linear equation system defining $\alpha(u)$, computing $\alpha_1(u) = 1 - \sum_{i\neq 1, i\in I_1} \alpha_i(u)$ from the second equation and rearranging terms in the first we obtain

$$\begin{aligned} &\bar{V}^\top \left(\sum_{i\neq 1, i\in I_1} \alpha_i(u) v_{ii}(\mathcal{X}(u) \right) + \sum_{(k,l)\in I_2} \alpha_j v_{kl}(\mathcal{X}(u))) \\ &= -\bar{V}^\top \left(\left(1 - \sum_{i\neq 1, i\in I_1} \alpha_i(u) \right) v_{11}(\mathcal{X}(u)) - \bar{g} \right). \end{aligned}$$

From both sides of the above formula subtracting $\bar{V}^\top \sum_{i\neq 1, i\in I_1} \alpha_i(u) v_{ii}(\mathcal{X}(u))$, and by the definition of $V(u)$ in Theorem 4.1 we get

$$\left(\bar{V}^\top V(u) \right) \left[\{\alpha_i(u)\}_{i\neq 1} \cup \{\alpha_j(u)\} \right] = -\bar{V}^\top \left(v_{11}\left(\mathcal{X}(u)\right) - \bar{g} \right),$$

and the conclusion follows.

In particular, for $u = 0$, because of $\mathcal{X}(0) = \bar{x}$, and $\alpha = \bar{\alpha}$ therein uniquely satisfies the linear system above; so $\alpha(0) = \bar{\alpha}$. □

Because both the gradients and the multipliers are C^1, it is possible to obtain a second-order expansion for λ_1, simply from the derivatives of these objects. Now we are in a position to give the specific expressions for the $\mathcal{U}$-Lagrangian and its derivatives.

Theorem 4.2 *Suppose the transversality condition (T) holds at $\bar{x}$, with the corresponding trajectory $\mathcal{X}(u)$ and basic matrix $\bar{V}$, and consider the $\mathcal{U}$-Lagrangian function defined in (2.2), and the multiplier functions $\alpha_i(u)$ from Proposition 4.1. Then for all u small enough, we have the following:*

(i) the vector $\bar{V}v(u)$ is an element of $w(u)$. Equivalently, the trajectory vector $\mathcal{X}(u)$ is an element of $\bar{x} + \bar{U}u + w(u)$;

(ii) the $\mathcal{U}$-Lagrangian $L_{\mathcal{U}}$ is given by

$$\begin{aligned} L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) &= \lambda_1(\mathcal{X}(u)) - \bar{g}^\top \bar{V} v(u) \\ &= q_i^\top(\mathcal{X}(u)) A(\mathcal{X}(u)) q_i(\mathcal{X}(u)) - \bar{g}^\top \bar{V} v(u) \end{aligned}$$

and

$$q_k^\top(\mathcal{X}(u)) A(\mathcal{X}(u)) q_l(\mathcal{X}(u)) = 0;$$

(iii) the gradient of $L_{\mathcal{U}}$ is

$$\nabla L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) = \bar{U}^\top g(u),$$

where $g(u)$ is defined by

$$g(u) = \sum_{i \in I_1} \alpha_i v_{ii}(\mathcal{X}(u)) + \sum_{(k,l) \in I_2} \alpha_j v_{kl}(\mathcal{X}(u));$$

(iv) the Hessian of $L_{\mathcal{U}}$ is given by

$$\nabla^2 L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) = J\mathcal{X}(u)^\top M(u) J\mathcal{X}(u),$$

where $M(u)$ is the $n \times n$ matrix function defined by

$$\begin{aligned} M(u) = &\sum_{i \in I_1} \alpha_i(u) \nabla^2 \left[q_i^\top(\mathcal{X}(u)) A(\mathcal{X}(u)) q_i(\mathcal{X}(u)) \right] \\ &+ \sum_{(k,l) \in I_2} \alpha_j(u) \nabla^2 \left[q_k^\top(\mathcal{X}(u)) A(\mathcal{X}(u)) q_l(\mathcal{X}(u)) \right]; \end{aligned}$$

(v) in particular, $L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) = \lambda_1 \circ A(\bar{x})$, and the $\mathcal{U}$-gradient and $\mathcal{U}$-Hessian at $\bar{x}$ are expressed as

$$\nabla L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) = \bar{U}^\top \bar{g} = \bar{U}^\top g, \quad for\ all\ g \in \partial \lambda_1 \circ A(\bar{x})$$

and

$$\nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) = \bar{U}^\top M(0) \bar{U}.$$

Proof (i) It is not difficult to obtain the result through the definition of $w(u)$ and the transversality condition.

(ii) Through the construction of $v(u)$ in Theorem 4.1 (i), and the transversality condition,

$$\begin{aligned}\lambda_1(\mathcal{X}(u)) &= \max\left\{q_i^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_i(\mathcal{X}(u))\right\}\\ &= q_i^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_i(\mathcal{X}(u)),\end{aligned}$$

so

$$\begin{aligned}L_{\mathcal{U}}(u;\bar{g}_{\mathcal{V}}) &= \lambda_1(\mathcal{X}(u)) - \bar{g}^\top\bar{V}v(u)\\ &= q_i^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_i(\mathcal{X}(u)) - \bar{g}^\top\bar{V}v(u).\end{aligned}$$

(iii) Making use of the chain rule to differentiate the following system with respect to u,

$$\begin{cases} L_{\mathcal{U}}(u;\bar{g}_{\mathcal{V}}) = q_i^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_i(\mathcal{X}(u)) - \bar{g}^\top\bar{V}v(u) \; i\in I_1\\ q_k^\top(\mathcal{X}(u))A(\mathcal{X}(u))q_l(\mathcal{X}(u)) = 0, \qquad (k,l)\in I_2,\end{cases}$$

we obtain

$$\begin{cases} \nabla L_{\mathcal{U}}(u;\bar{g}_{\mathcal{V}}) = J\mathcal{X}(u)^\top v_{ii}(\mathcal{X}(u)) - Jv(u)^\top\bar{V}^\top\bar{g},\; i\in I_1\\ J\mathcal{X}(u)^\top v_{kl}(\mathcal{X}(u)) = 0, \qquad (k,l)\in I_2.\end{cases}$$

On both sides of the above equation system, multiplying by the corresponding multipliers $\alpha_i(u)$ and $\alpha_j(u)$, respectively, adding these results, and using the fact that $\sum_{i\in I_1}\alpha_i(u) = 1$ provides

$$\nabla L_{\mathcal{U}}(u;\bar{g}_{\mathcal{V}}) = J\mathcal{X}(u)^\top g(u) - Jv(u)^\top\bar{V}^\top\bar{g},$$

where $g(u) = \sum_{i\in I_1}\alpha_i v_{ii}(\mathcal{X}(u)) + \sum_{(k,l)\in I_2}\alpha_j v_{kl}(\mathcal{X}(u))$. Through the transpose of the expression of $J\mathcal{X}(u)$, we gain

$$\nabla L_{\mathcal{U}}(u;\bar{g}_{\mathcal{V}}) = \bar{U}^\top g(u) - Jv(u)^\top\bar{V}^\top(g(u)-\bar{g}),$$

which combining the fact that $\bar{V}^\top(g(u)-\bar{g}) = 0$ gives the desired result.

(iv) Differentiating (iii) with respect to u, we get

$$\begin{aligned}\nabla^2 L_{\mathcal{U}}(u;\bar{g}_{\mathcal{V}}) = \bar{U}^\top M(u)J\mathcal{X}(u) + \bar{U}^\top &\left[\sum_{i\in I_1}\alpha_i(u)v_{ii}(\mathcal{X}(u))J\alpha_i(u)\right.\\ &\left.+\sum_{(k,l)\in I_2}\alpha_j(u)v_{kl}(\mathcal{X}(u))J\alpha_j(u)\right],\end{aligned}$$

where

$$M(u) = \sum_{i \in I_1} \alpha_i(u) \nabla^2 \left[q_i^\top (\mathcal{X}(u)) A(\mathcal{X}(u)) q_i(\mathcal{X}(u)) \right] + \sum_{(k,l) \in I_2} \alpha_j(u) \nabla^2 \left[q_k^\top (\mathcal{X}(u)) A(\mathcal{X}(u)) q_l(\mathcal{X}(u)) \right].$$

It follows that

$$\sum_{i \in I_1} \alpha_i(u) v_{ii}(\mathcal{X}(u)) J\alpha_i(u) + \sum_{(k,l) \in I_2} \alpha_j(u) v_{kl}(\mathcal{X}(u)) J\alpha_j(u) = -V(u)(\bar{V} V(u))^{-1} \bar{V}^\top M(u) J\mathcal{X}(u).$$

Then

$$\begin{aligned} \nabla^2 L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) &= \bar{U}^\top M(u) J\mathcal{X}(u) - \bar{U}^\top V(u)(\bar{V} V(u))^{-1} \bar{V}^\top M(u) J\mathcal{X}(u) \\ &= \bar{U}^\top M(u) J\mathcal{X}(u) + Jv(u)^\top \bar{V}^\top M(u) J\mathcal{X}(u) \\ &= [\bar{U}^\top + Jv(u)^\top \bar{V}^\top] M(u) J\mathcal{X}(u) \\ &= J\mathcal{X}(u)^\top M(u) J\mathcal{X}(u). \end{aligned}$$

(v) Because $v(0) = 0$, $\mathcal{X}(0) = \bar{x}$, $\mathcal{X}(0) = \bar{U}$, $\sum_{i \in I_1} \alpha_i(0) = 1$,

$$\begin{aligned} L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) &= \lambda_1 \circ A(\bar{x}) - \langle \bar{g}_{\mathcal{V}}, v(0) \rangle = \lambda_1 \circ A(\bar{x}) - \langle \bar{g}_{\mathcal{V}}, 0 \rangle = \lambda_1 \circ A(\bar{x}) \\ \nabla L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) &= \bar{U}^\top g(0) = \bar{U}^\top \bar{g} \\ \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) &= J\mathcal{X}(0)^\top M(0) J\mathcal{X}(0) = \bar{U}^\top M(0) \bar{U}. \end{aligned}$$

□

The maximum eigenvalue function is structural and qualified enough that a second-order object "$\mathcal{U}$-Hessian" exists. Moreover, this allows for the construction of the smooth trajectories, tangent to $\mathcal{U}$, along which the function is twice differentiable. We have the following theorem:

Theorem 4.3 *Suppose the transversality condition (T) holds, with the corresponding smooth trajectory $\mathcal{X}(u) = \bar{x} + \bar{U}u + \bar{V}v(u)$. Then for each u small enough, we have the second-order expansion*

$$\lambda_1(\mathcal{X}(u)) = \lambda_1 \circ A(\bar{x}) + \bar{g}_{\mathcal{U}}^\top u + \bar{g}^\top \bar{V} v(u) + \frac{1}{2} u^\top \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) u + o(\|u\|^2). \quad (4.5)$$

This can also be written as

$$\lambda_1(\mathcal{X}(u)) = \lambda_1 \circ A(\bar{x}) + \bar{g}^\top (\mathcal{X}(u) - \bar{x}) + \frac{1}{2} (\mathcal{X}(u) - \bar{x})^\top \bar{U} \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) \bar{U} (\mathcal{X}(u) - \bar{x}) + o(\|(\mathcal{X}(u) - \bar{x})\|^2).$$

Proof Since $L_{\mathcal{U}} \in C^2$, we obtain

$$\begin{aligned} L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) &= L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) + \langle \nabla L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}), u \rangle_{\mathcal{U}} + \tfrac{1}{2} u^\top \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) u + o(\|u\|^2) \\ &= \lambda_1 \circ A(\bar{x}) + \langle \bar{g}_{\mathcal{U}}, u \rangle_{\mathcal{U}} + \tfrac{1}{2} u^\top \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) u + o(\|u\|^2). \end{aligned} \tag{4.6}$$

From the definition of $L_{\mathcal{U}}$, we have

$$L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) = \lambda_1(\mathcal{X}(u)) - \langle \bar{g}_{\mathcal{V}}, v(u) \rangle_{\mathcal{V}}.$$

Therefore,

$$\begin{aligned} \lambda_1(\mathcal{X}(u)) &= L_{\mathcal{U}}(u; \bar{g}_{\mathcal{V}}) + \langle \bar{g}_{\mathcal{V}}, v(u) \rangle_{\mathcal{V}} \\ &= \lambda_1 \circ A(\bar{x}) + \langle \bar{g}_{\mathcal{U}}, u \rangle_{\mathcal{V}} + \langle \bar{g}_{\mathcal{V}}, v(u) \rangle_{\mathcal{V}} + \frac{1}{2} u^\top \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) u + o(\|u\|^2) \\ &= \lambda_1 \circ A(\bar{x}) + \langle \bar{g}, u \oplus v(u) \rangle + \frac{1}{2} u^\top \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) u + o(\|u\|^2) \\ &= \lambda_1 \circ A(\bar{x}) + \langle \mathcal{X}(u) - \bar{x} \rangle + \frac{1}{2} (\mathcal{X}(u) - \bar{x})^\top \bar{U} \nabla^2 L_{\mathcal{U}}(0; \bar{g}_{\mathcal{V}}) \\ &\quad \cdot \bar{U}^\top (\mathcal{X}(u) - \bar{x}) + o(\|\mathcal{X}(u) - \bar{x}\|^2). \end{aligned}$$

The proof is done. □

Second-order $\mathcal{U}$-derivatives are useful for specifying the second-order expansions for λ_1, as is shown above. Meanwhile, it also gives the related optimality condition for optimizers, as is shown in §4.3. We call the corresponding Hessian of $L_{\mathcal{U}}$ at $u = 0$ a $\mathcal{U}$-Hessian for λ_1 at $\bar{x}$ and denote it by $\bar{H} := \nabla^2 L_{\mathcal{U}}(0; 0)$.

4.2 Decomposition algorithm

Suppose $0 \in \partial \lambda_1 \circ A(\bar{x})$, next we present a conceptual $\mathcal{VU}$ algorithm framework that can solve (P). This algorithm performs a minimizing step in the $\mathcal{V}$ subspace, following by a $\mathcal{U}$-Newton step in order to get the superlinear convergence.

Algorithm 4.1

Step 0 Initialization. Given $\epsilon > 0$. Choose a starting point x^0 close to $\bar{x}$, and a subgradient $\tilde{g}^0 \in \partial \lambda_1 \circ A(x^0)$, set $\Bbbk = 0$.

Step 1 Stopping test. Stop if

$$\|\tilde{g}^{\Bbbk}\|^2 \leq \epsilon.$$

Step 2 Construct the $\mathcal{VU}$ decomposition at $\bar{x}$, i.e., $R^m = \mathcal{U}(\bar{x}) \oplus \mathcal{V}(\bar{x})$. Compute

$$\nabla^2 L_{\mathcal{U}}(0; 0) = \bar{U}^\top M(0) \bar{U},$$

where

$$M(0) = \sum_{i \in I_1} \bar{\alpha}_i \nabla^2 \left[q_i^\top(\bar{x}) A(\bar{x}) q_i(\bar{x}) \right] + \sum_{(\Bbbk, l) \in I_2} \bar{\alpha}_j \nabla^2 \left[q_{\Bbbk}^\top(\bar{x}) A(\bar{x}) q_l(\bar{x}) \right].$$

Step 3 Perform $\mathcal{V}$-step. Compute

$$\delta_{\mathcal{V}}^{\Bbbk} \in \arg\min\left\{\lambda_1\left(x^{\Bbbk} + 0 \oplus \delta_{\mathcal{V}}\right) : \delta_{\mathcal{V}} \in \mathcal{V}\right\}.$$

Set $\tilde{x}^{\Bbbk} = x^{\Bbbk} + 0 \oplus \delta_{\mathcal{V}}^{\Bbbk}$.

Step 4 Perform $\mathcal{U}$-step. Compute $\delta_{\mathcal{U}}^{\Bbbk}$ from the system

$$\bar{U}^{\top} M(0) \bar{U} \delta_{\mathcal{U}} + \bar{U}^{\top} \tilde{g}^{\Bbbk} = 0, \tag{4.7}$$

where

$$\sum_{i \in I_1} \alpha_i(u) v_{ii}(\tilde{x}^{\Bbbk}) + \sum_{(\Bbbk,l) \in I_2} \alpha_j(u) v_{\Bbbk l}(\tilde{x}^{\Bbbk}) = \tilde{g}^{\Bbbk} \in \partial \lambda_1 \circ A(\tilde{x}^{\Bbbk})$$

is such that $\bar{V}^{\top} \tilde{g}^{\Bbbk} = 0$. Compute $x^{\Bbbk+1} = \tilde{x}^{\Bbbk} + \delta_{\mathcal{U}}^{\Bbbk} \oplus 0 = x^{\Bbbk} + \delta_{\mathcal{U}}^{\Bbbk} \oplus \delta_{\mathcal{V}}^{\Bbbk}$.

Step 5 Update. Set $\Bbbk = \Bbbk + 1$, and return to Step 1.

Next we shall discuss the convergence of Algorithm 4.1.

Theorem 4.4 *Suppose* $0 \in \mathrm{ri}\,\partial \lambda_1 \circ A(\bar{x})$, $\nabla^2 L_{\mathcal{U}}(0; 0) \succ 0$. *Then the iterate points* $\{x^{\Bbbk}\}_{\Bbbk=1}^{\infty}$ *generated by Algorithm 4.1 converge and satisfy*

$$\|x^{\Bbbk+1} - \bar{x}\| = o(\|x^{\Bbbk} - \bar{x}\|).$$

Proof Set $u^{\Bbbk} = (x^{\Bbbk} - \bar{x})_{\mathcal{U}}$, $v^{\Bbbk} = (x^{\Bbbk} - \bar{x})_{\mathcal{V}} + \delta_{\mathcal{V}}^{\Bbbk}$. Then $\bar{x} + u^{\Bbbk} \oplus v^{\Bbbk} = x^{\Bbbk} + 0 \oplus \delta_{\mathcal{V}}$, and

$$\begin{aligned} \delta_{\mathcal{V}}^{\Bbbk} &\in \arg\min\left\{\lambda_1(x^{\Bbbk} + 0 \oplus \delta_{\mathcal{V}}) : \delta_{\mathcal{V}} \in \mathcal{V}\right\} \\ &= \arg\min\left\{\lambda_1(\bar{x} + u^{\Bbbk} \oplus v^{\Bbbk}) : \delta_{\mathcal{V}} \in \mathcal{V}\right\}. \end{aligned}$$

Hence, $v^{\Bbbk} \in W(u^{\Bbbk}; 0)$. It follows that

$$\left\|\left(x^{\Bbbk+1} - \bar{x}\right)_{\mathcal{V}}\right\| = \left\|\left(\tilde{x}^{\Bbbk} - \bar{x}\right)_{\mathcal{V}}\right\| = o\left(\left\|\left(x^{\Bbbk} - \bar{x}\right)_{\mathcal{U}}\right\|\right) = o\left(\left\|\left(x^{\Bbbk} - \bar{x}\right)\right\|\right). \tag{4.8}$$

Since $\nabla^2 L_{\mathcal{U}}(0; 0)$ exists and $\nabla L_{\mathcal{U}}(0; 0) = 0$, we have

$$\nabla L_{\mathcal{U}}(u^{\Bbbk}; 0) = \bar{U}^{\top} \tilde{g}^{\Bbbk} = 0 + \nabla^2 L_{\mathcal{U}}(0; 0) u^{\Bbbk} + o(\|u^{\Bbbk}\|_{\mathcal{U}}).$$

Because of (4.7), we have $\nabla^2 L_{\mathcal{U}}(0; 0)(u^{\Bbbk} + \delta_{\mathcal{U}}^{\Bbbk}) = o(\|u^{\Bbbk}\|_{\mathcal{U}})$. From the assumption $\nabla^2 L_{\mathcal{U}}(0; 0) \succ 0$ it follows that $\nabla^2 L_{\mathcal{U}}(0; 0)$ is invertible and $\|u^{\Bbbk} + \delta_{\mathcal{U}}^{\Bbbk}\| = o(\|u^{\Bbbk}\|_{\mathcal{U}})$. As a result, one gets

$$\begin{aligned}&(x^{\mathbb{k}+1}-\bar{x})_{\mathcal{U}}\\&=(x^{\mathbb{k}+1}-\tilde{x}^{\mathbb{k}})_{\mathcal{U}}+(\tilde{x}^{\mathbb{k}}-x^{\mathbb{k}})_{\mathcal{U}}+(x^{\mathbb{k}}-\bar{x})_{\mathcal{U}}\\&=u^{\mathbb{k}}+\delta^{\mathbb{k}}_{\mathcal{U}}.\end{aligned}$$

Then

$$\left\|\left(x^{\mathbb{k}+1}-\bar{x}\right)_{\mathcal{U}}\right\|=o\left(\left\|u^{\mathbb{k}}\right\|_{\mathcal{U}}\right)=o\left(\left\|x^{\mathbb{k}}-\bar{x}\right\|\right). \tag{4.9}$$

Combining (4.8) and (4.9), the proof is complete. □

5 Application

In this section, we mainly study the explicit $\mathcal{VU}$-decomposition forms of matrix convexity from those examples in Sect. 2.2. Our purpose here is to demonstrate that $\mathcal{U}$-Lagrangian theory can potentially be very efficient in solving some actual problems. With regard to their proofs, we omit all unnecessary details for brevity and give some important conclusions. We note that though the ideas about the results are similar, as the readers will find, the technical details become much more involved. Yet, their formulation is simple, chain rules are not easy to obtain, we still need to introduce some geometrical conditions to get them.

Minimizing a linear objective function subject to bilinear matrix inequality (BMI) constraints is a useful way to describe many robust control synthesis problems. Many other problems in automatic control lead to BMI feasibility and optimization programs, such as filtering problems, synthesis of structured or reduced-order feedback controllers, robustness, simultaneous stabilization problems and many others (Fukuda and Kojima 2001; Thevenet et al. 2006). Formally, these problems may be described as

$$\begin{cases}\min\limits_{x\in R^m} & c^{\top}x\\[2ex] \text{s.t.} & \mathcal{B}(x)\preceq 0,\end{cases} \tag{5.1}$$

where the bilinear operator $\mathcal{B}: R^m \to S_n$ maps R^m into the space S_n of symmetric $n\times n$ matrices, $mn\times mn$ block matrix $[B_{i,j}]_{i,j=1}^m$ is positive semidefinite,

$$\mathcal{B}(x) := A_0+\sum_{i=1}^{m}x_iA_i+\sum_{i,j=1}^{m}x_ix_jB_{i,j}, \tag{5.2}$$

we call (5.1) a bilinear matrix inequality (BMI) problem. During the past decade, the importance of (BMI) for applications in automatic control has been recognized. The problem of the form (5.1) is clearly related to eigenvalue optimization. We consider the unconstrained problem

$$\min\ \ \lambda_1(\mathcal{B}(x)),\ \ x\in R^m, \tag{5.3}$$

and the constrained eigenvalue optimization problem

$$\begin{cases} \min\limits_{x \in R^m} \ c^\top x \\ \text{s.t.} \ \ \lambda_1(\mathcal{B}(x)) \leq 0. \end{cases} \tag{5.4}$$

The convexity of (5.3) and (5.4) is induced by the operator $\mathcal{B}$. Clearly, (5.1) is equivalent to (5.4). Problems of the form (5.4) may be transformed into (5.3) via exact penalization, even though it may be preferable to use the structure of (5.4) explicitly. We denote the exact penalty function

$$\mathcal{P}(x, \mu) := c^\top x + \mu[\lambda_1(\mathcal{B}(x))]^+ = c^\top x + \mu\lambda_1[\text{diag}(0, \mathcal{B}(x))],$$

where the penalty parameter $\mu \geq \mu^*$ for some threshold value μ^*, the sign diag stands for the diagonalized operation. We give the following assumption:

Assumption 5.1 (*Slater Constraint Qualification*) Assume there exists at least one x such that $\lambda_1(\mathcal{B}(x)) < 0$.

In view of Assumption 5.1, we will present the existence of the exact penalty function and estimate the upper bound of the penalty parameter μ by recapitulating Proposition 7.3.1 in Bertsekas et al. (2003) and Corollary 2.2 in Mangasarian (1985). We have the following

Proposition 5.1 (Exact penalty characterization of solvable convex programs) *Let λ_1 be defined as the form* (5.3)*, suppose the Slater constraint qualification (5.1) holds at some point x^1 and $\bar{x}$ is a solution of* (5.4)*. A necessary and sufficient condition for $\tilde{x} \in R^m$ to solve the minimization problem* (5.4) *is that $\tilde{x}$ minimizes $\mathcal{P}(x, \mu)$ over x in R^m for each $\mu \geq \hat{\mu}$, where $\hat{\mu} \in R_+$ is any dual optimal multiplier for* (5.4)*. Moreover, we have*

$$\hat{\mu} \leq \mu^* := \frac{c^\top(x^1 - \bar{x})}{-\lambda_1(\mathcal{B}(x^1))}.$$

Therefore, convex BMI problem (5.1) can be replaced by the following nonsmooth convex eigenvalue optimization problem

$$\min_{x \in R^m} \mathcal{P}(x, \mu) = c^\top x + \mu\lambda_1(\mathcal{H}(x)), \tag{5.5}$$

where the mapping $\mathcal{H} : R^m \to S_{n+1}$ is defined by $\mathcal{H}(x) = \text{diag}(0, \mathcal{B}(x)) = \begin{pmatrix} 0 & \\ & \mathcal{B}(x) \end{pmatrix}$, and the penalty parameter $\mu \geq \mu^*$ for some threshold value μ^*.

Remark 5.1 In some cases, for example, when the diagonal of $\mathcal{B}(x)$ is fixed or free, transforming (5.4) to (5.5) can be done efficiently because an efficient value for the penalty parameter is known. Such is often the case when (5.4) is the relaxation of a graph problem; see Goemans (1997) or Grötschel et al. (1983) for more on the origins of these problems.

Note that because $\mathcal{B}(x)$ is differentiable, then its differential at a point $x \in R^m$ can be written in the form $D\mathcal{B}(x)h = \sum_{i=1}^{m} h_i B_i(x)$, where $B_i(x) := \partial\mathcal{B}(x)/\partial x_i = A_i + 2\sum_{j=1}^{m} x_j B_{ij}$, $i = 1, \ldots, m$, are $n \times n$ matrices of partial derivatives of the elements of $G(x)$ with respect to x_i; therefore, we have $D\mathcal{H}(x)h = \begin{pmatrix} 0 & \\ & D\mathcal{B}(x) \end{pmatrix} h = \begin{pmatrix} 0 & \\ & D\mathcal{B}(x)h \end{pmatrix} = \begin{pmatrix} 0 & \\ & \sum\limits_{i=1}^{m} h_i B_i(x) \end{pmatrix}$, and $B_i(x)$ is defined in above form. The adjoint operator of the derivative $D\mathcal{B}(x)^\star : S_n \to R^m$ is given by the following formula

$$\begin{aligned} D\mathcal{B}(x)^\star\Omega &= \left(\langle \tfrac{\partial\mathcal{B}(x)}{\partial x_1}, \Omega\rangle, \cdots, \langle \tfrac{\partial\mathcal{B}(x)}{\partial x_m}, \Omega\rangle\right)^\top \\ &= \left(\langle A_1 + 2\sum_{j=1}^{m} x_j B_{1j}, \Omega\rangle, \cdots, \langle A_m + 2\sum_{j=1}^{m} x_j B_{mj}, \Omega\rangle\right)^\top, \quad \Omega \in S_n, \end{aligned}$$

and $D\mathcal{H}(x)^\star\Xi = D\mathcal{B}(x)^\star\Xi_4 = \left(\langle A_1 + 2\sum\limits_{j=1}^{m} x_j B_{1j}, \Xi_4\rangle, \cdots, \langle A_m + 2\sum\limits_{j=1}^{m} x_j B_{mj}, \Xi_4\rangle\right)^\top$, $\Xi = \begin{pmatrix} \Xi_1 & \Xi_2 \\ \Xi_3 & \Xi_4 \end{pmatrix} \in S_{n+1}$, $\Xi_1 \in R^1$, $\Xi_2 \in R^{1\times n}$, $\Xi_3 \in R^{n\times 1}$, $\Xi_4 \in S_n$.

Firstly, we can easily obtain a well-known description of the subdifferential about $P(x, \mu)$ in the sense of convex analysis.

Proposition 5.2

$$\partial\mathcal{P}(x, \mu) = c + \mu D\mathcal{H}(x)^\star\partial\lambda_1(\mathcal{H}(x)), \tag{5.6}$$

where $\partial\lambda_1(\mathcal{H}(x)) = \{U(x) \in S_{n+1} : Z \in S_r^+, \mathrm{tr}Z = 1, U(x) = Q_1(x)ZQ_1(x)^\top\}$, *using the above form of* $D\mathcal{H}(x)^\star$,

$$\begin{aligned} \partial\mathcal{P}(x, \mu) &= \{c + \mu D\mathcal{H}(x)^\star U(x), U(x) \in \partial\lambda_1(\mathcal{H}(x))\} \\ &= \{c + \mu D\mathcal{B}(x)^\star[U(x)]_4, U(x) \in \partial\lambda_1(\mathcal{H}(x))\} \\ &= \left\{ c + \mu\left(\langle A_1 + 2\sum_{j=1}^{m} x_j B_{1j}, [U(x)]_4\rangle, \cdots, \langle A_m + 2\sum_{j=1}^{m} x_j B_{mj}, [U(x)]_4\rangle\right)^\top, \right. \\ &\qquad \left. U(x) \in \partial\lambda_1(\mathcal{H}(x)) \right\}, \end{aligned}$$

where $U(x) = \begin{pmatrix} [U(x)]_1 & [U(x)]_2 \\ [U(x)]_3 & [U(x)]_4 \end{pmatrix} \in S_{n+1}$, $[U(x)]_1 \in R^1$, $[U(x)]_2 \in R^{1\times n}$, $[U(x)]_3 \in R^{n\times 1}$, $[U(x)]_4 \in S_n$.

The relative interior of $\partial(\lambda_1(\mathcal{H}(x)))$ *has the expression*

$$\mathrm{ri}\partial\mathcal{P}(x, \mu) = c + \mu D\mathcal{H}(x)^\star\mathrm{ri}\partial\lambda_1(\mathcal{H}(x)), \tag{5.7}$$

where $\mathrm{ri}\partial\lambda_1(\mathcal{H}(x)) = \{U(x) \in S_{n+1} : Z \in S_r, Z \succ 0, \mathrm{tr}Z = 1, U(x) = Q_1(x)ZQ_1(x)^\top\}$.

If $r = 1$, *then* $\mathcal{P}(x, \mu)$ *is differentiable, and* $\nabla\mathcal{P}(x, \mu) = c + \mu D\mathcal{H}(x)^\star(Q_1(x) Q_1(x)^\top)$.

Next we might assume that $D\mathcal{H}(x)$ at x^* is onto, so $D\mathcal{H}(x^*)$ is invertible.

We have the following theorems. Since these results can be obtained in a similar way to those in Sect.3, the proofs in detail will be omitted.

Theorem 5.1 *(1) The subspaces* $\mathcal{U}_{\mathcal{P}_\mu}(x^*)$ *and* $\mathcal{V}_{\mathcal{P}_\mu}(x^*)$ *are respectively characterized by*

$$\begin{aligned}\mathcal{U}_{\mathcal{P}_\mu}(x^*) &= D\mathcal{H}(x^*)^{-1}\mathcal{U}_{\lambda_1}(\mathcal{H}(x^*)) \\ &= D\mathcal{H}(x^*)^{-1}\Big\{U \in S_n : Q_1(x^*)^\top U Q_1(x^*) \\ &\quad -\frac{1}{r}\mathrm{tr}(Q_1(x^*)^\top U Q_1(x^*))I_r = 0\Big\}\end{aligned} \tag{5.8}$$

and

$$\begin{aligned}\mathcal{V}_{\mathcal{P}_\mu}(x^*) &= D\mathcal{H}(x^*)^\star\mathcal{V}_{\lambda_1}(\mathcal{H}(x^*)) \\ &= D\mathcal{H}(x^*)^\star\left\{Q_1(x^*)ZQ_1(x^*)^\top : Z \in S_r^+, \mathrm{tr}Z = 0\right\}.\end{aligned} \tag{5.9}$$

(2) If $r = 1$, *then* $\mathcal{P}(x^*, \mu)$ *is a differentiable function, here we have* $\mathcal{U}_{\mathcal{P}_\mu}(x^*) = R^m$, *and* $\mathcal{V}_{\mathcal{P}_\mu}(x^*) = \{0\}$.

(3) If $\mathrm{int}\ \partial\mathcal{P}(x^*, \mu) \neq \emptyset$, *then we have* $\mathcal{U}_{\mathcal{P}_\mu}(x^*) = \{0\}$, *and* $\mathcal{V}_{\mathcal{P}_\mu}(x^*) = R^m$.

Take $g^* \in \mathrm{ri}\partial\mathcal{P}(x^*, \mu)$ and define the $\mathcal{U}$-Lagrangian of $\mathcal{P}_\mu(x^*) := \mathcal{P}(x^*, \mu)$ at (x^*, g^*) according to (2.1); in the following we denote it by $L_{\mathcal{U},\mathcal{P}_\mu}(x^*, g^*; \cdot)$. From Theorem 2.1, $L_{\mathcal{U},\mathcal{P}_\mu}(x^*, g^*; \cdot)$ is differentiable at $u = 0$. We obtain the following composition rule.

Theorem 5.2 *Let* $G^* \in \mathrm{ri}\partial\lambda_1(\mathcal{H}(x^*))$ *be such that* $g^* = c + \mu D\mathcal{H}(x^*)^\star \cdot G^*$. *Then,*

$$\begin{aligned}\nabla L_{\mathcal{U},\mathcal{P}_\mu}(x^*, g^*; 0) &= \mathrm{proj}_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)}c \\ &+ \left[\mathrm{proj}_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)} \circ D\mathcal{H}(x^*)^\star \circ \mathrm{proj}^*_{\mathcal{U}_{\lambda_1}(\mathcal{H}(x^*))}\right] \cdot \nabla L_{\mathcal{U},\lambda_1}\left(\mathcal{H}(x^*), G^*; 0\right),\end{aligned} \tag{5.10}$$

where $\mathcal{U}_{\mathcal{P}_\mu}(x^*)$ *is given by (5.8).*

We then introduce the subspace

$$\mathcal{T} := \{Z \in S_r : \mathrm{tr}Z = 0\}, \tag{5.11}$$

and consider the mapping

$$\Phi : B(\mathcal{H}^*, \delta) \ni \mathcal{H} \mapsto Q_{tot}(\mathcal{H})^\top \mathcal{H} Q_{tot}(\mathcal{H}) - \frac{1}{r}\mathrm{tr}\left(Q_{tot}(\mathcal{H})^\top \mathcal{H} Q_{tot}(\mathcal{H})\right) I_r \in \mathcal{T}. \tag{5.12}$$

As pointed out by M. Overton through several works (Overton 1992; Overton and Womersley 1995), the maximum eigenvalue function enjoys a very specific structure in a neighborhood of a point $Y \in S_{n+1}$ belonging to the set $\mathcal{M}_r$, the function λ_1 is smooth on $\mathcal{M}_r$. More explicitly, the set $\mathcal{M}_r$ is a submanifold of S_{n+1} and there exists a neighborhood Ω of Y in S_{n+1} such that the function $\mathcal{M}_r \cap \Omega \ni X \mapsto \lambda_1(X)$ is C^∞. To obtain a similar result for the function $\mathcal{P}(x^*, \mu)$, some precautions must be taken: the intersection of $\mathcal{M}_r$ with the image of the mapping $R^m \ni x \mapsto \mathcal{H}(x)$ may have some singularities. To avoid these situations, a transversality assumption can be made:

Definition 5.1 We say that the mapping $\mathcal{H}(\cdot)$ is transversal to $\mathcal{M}_r$ at $x^* \in \mathcal{H}^{-1}\mathcal{M}_r$ if and only if

$$(\mathbf{T}') \quad range D\mathcal{H}(x^*) + \mathcal{U}_{\lambda_1}(\mathcal{H}(x^*)) = S_{n+1}. \tag{5.13}$$

Next we obtain a local equation of $\mathcal{W}_r := \mathcal{H}^{-1}\mathcal{M}_r$ via a simple composition rule, which plays an important role in our subsequent analysis.

Theorem 5.3 *If ($\mathbf{T}'$) is satisfied at x^*, then there exists $\rho > 0$ such that $\varphi(x) = 0$, where $\varphi : B(x^*, \rho) \ni x \mapsto \Phi(\mathcal{H}(x)) \in S_r$ and Φ is given by (5.12), is a local equation of $\mathcal{W}_r \cap B(x^*, \rho)$. Moreover, for all $x \in B(x^*, \rho)$, we have*

$$\mathrm{T}_{\mathcal{W}_r}(x) = \ker D\varphi(x).$$

Theorem 5.4 *Assume ($\mathbf{T}'$) is satisfied at x^* and take $g^* \in \mathrm{ri}\partial\mathcal{P}(x^*, \mu)$. Then*

(1) the subspaces $\mathcal{U}_{\mathcal{P}_\mu}(x^*)$ and $\mathcal{V}_{\mathcal{P}_\mu}(x^*)$ are respectively the tangent and normal spaces to $\mathcal{W}_r$ at x^*.
(2) there exists $\rho > 0$ and a C^∞-mapping $v : \mathcal{U}_{\mathcal{P}_\mu}(x^*) \cap B(0, \rho) \to \mathcal{V}_{\mathcal{P}_\mu}(x^*)$ such that the mapping

$$p_{x^*} : \mathcal{U}_{\mathcal{P}_\mu}(x^*) \cap B(0, \rho) \ni u \mapsto x^* + u \oplus v(u), \tag{5.14}$$

is a tangential parametrization of $\mathcal{W}_r$.

Theorem 5.5 *Assume (T') is satisfied at x^* and take $g^* \in \mathrm{ri}\,\partial\mathcal{P}(x^*, \mu)$. Then there exists $\eta > 0$ such that for all $u \in B(0, \rho) \subset \mathcal{U}_{\mathcal{P}_\mu}(x^*)$, the set $w(u)$ is a singleton:*

$$w(u) = \{v(u)\} \text{ for all } u \in B(0, \eta),$$

where $v(\cdot)$ is the C^∞-mapping defined in Theorem 5.4 (2).

Transversality is a sufficient condition to obtain the differentiability of the $\mathcal{U}$-Lagrangian of $\mathcal{P}_\mu$ and to compute its Hessian via simple chain rules. So this leads us to our main result.

Theorem 5.6 *Assume (T') is satisfied at x^* and take $g^* \in \operatorname{ri} \partial \mathcal{P}_\mu(x^*)$. Then the $\mathcal{U}$-Lagrangian function $L_{\mathcal{U},\mathcal{P}_\mu}(x^*, g^*; \cdot)$ is C^∞. Moreover, at $u = 0$,*

$$\nabla^2 L_{\mathcal{U},\mathcal{P}_\mu}(x^*, g^*; 0) = \operatorname{proj}_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)} \circ D\mathcal{H}(x^*)^\star \circ H(\mathcal{H}(x^*), G^*) \circ D\mathcal{H}(x^*) \circ \operatorname{proj}^*_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)}, \tag{5.15}$$

where G^* is the unique subgradient of $\partial \lambda_1(\mathcal{H}(x^*))$ such that $g^* = c + \mu D\mathcal{H}(x^*)^\star G^*$ and the operator $H(\mathcal{H}^*, G^*)$ is the symmetric operator defined by

$$H(\mathcal{H}^*, G^*) \cdot Y = G^* Y [\lambda_1^* I_{n+1} - \mathcal{H}^*]^\dagger + [\lambda_1^* I_{n+1} - \mathcal{H}^*]^\dagger Y G^*, \tag{5.16}$$

This can also be written

$$\nabla^2 L_{\mathcal{U},\mathcal{P}_\mu}(x^*, g^*; 0) = B(x^*)^* \circ \nabla^2 L_{\mathcal{U},\lambda_1}(\mathcal{H}(x^*), G^*; 0) \circ B(x^*), \tag{5.17}$$

where $B(x^*) = \operatorname{proj}_{\mathcal{U}_{\lambda_1}(\mathcal{H}(x^*))} \circ D\mathcal{H}(x^*) \circ \operatorname{proj}^*_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)}$ and $\mathcal{U}_{\mathcal{P}_\mu}(x^*)$ is given by (5.8).

Under the transversality condition, we obtain that the function $\mathcal{P}(x, \mu)$ is locally C^∞ on $\mathcal{W}_r$. We say that such a point is *regular*. We are now in a position to derive the second-order development of $\mathcal{P}_\mu$ along $\mathcal{W}_r$ at x^*.

Corollary 5.1 *If the condition of Theorem 5.6 is satisfied, we have for $x^* + d \in \mathcal{W}_r$ and $d \to 0$,*

$$\begin{aligned}\mathcal{P}_\mu(x^* + d) = \mathcal{P}_\mu(x^*) &+ \langle g^*, d\rangle \\ &+ \frac{1}{2}\left\langle \operatorname{proj}_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)} d, \nabla^2 L_{\mathcal{U},\mathcal{P}_\mu}(x^*, g^*; 0) \cdot \left(\operatorname{proj}_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)} d\right)\right\rangle_{\mathcal{U}_{\mathcal{P}_\mu}(x^*)} \\ &+ o(\|d\|^2).\end{aligned} \tag{5.18}$$

The theory developed so far strongly suggests the following algorithmic application: near a solution of (5.3), minimize the second-order development of the $\mathcal{U}$-Lagrangian of $\mathcal{P}_\mu$. Here we present a conceptual algorithm which relies on this simple idea.

Let us consider a minimum point x^* and call r the multiplicity of $\lambda_1(\mathcal{H}(x^*))$. Given $x \in B(x^*, \rho)$ for some $\rho > 0$, we need to compute some x_+, superlinearly closer to x^*. Consider the following conceptual algorithm.

Algorithm 5.1
$\mathcal{V}$-Step. Compute $\hat{x} \in \mathcal{W}_r$ a solution of

$$\min\{\|\hat{x} - x\| : \hat{x} \in \mathcal{W}_r\} \tag{5.19}$$

Dual-Step. Compute

$$g(\hat{x}) := \mathrm{proj}_{\partial \mathcal{P}_\mu(\hat{x})}(0) \tag{5.20}$$

$\mathcal{U}$-Step. Solve

$$\min_{u \in \mathcal{U}(\hat{x})} \left\langle u, \nabla L_{\mathcal{U},\mathcal{P}_\mu}(\hat{x}, g(\hat{x}); 0)\right\rangle_{\mathcal{U}_{\mathcal{P}_\mu}(\hat{x})} + \frac{1}{2}\left\langle u, \nabla^2 L_{\mathcal{U},\mathcal{P}_\mu}(\hat{x}, g(\hat{x}); 0)u\right\rangle_{\mathcal{U}_{\mathcal{P}_\mu}(\hat{x})}. \tag{5.21}$$

Update. Set $x_+ = \hat{x} + u$.

Remark 5.2 The above fast $\mathcal{VU}$-algorithm, which can also be called the predictor-corrector methods, follows a common two-step form. In our algorithm we have an iterate which lies on the active manifold, we use the smoothness of the function along the manifold to take a predictor step ($\mathcal{U}$-Newton step) in a direction tangent to the active manifold. Since, in general, the operation results in a point outside the manifold, it is followed by a corrector step ($\mathcal{V}$ step) which returns the iterate to the active manifold. Algorithms of this form can be found in Lemaréchal et al. (2000) and Mifflin and Sagastizábal (2005).

In order to get the quadratic convergence, we need to present the definition of strict complementarity and non-degeneracy condition, which can be seen as the generalization of the regularity assumption needed in all Newton-type methods.

Definition 5.2 (*Strict Complementarity*) We say that the strict complementarity *(SC)* holds at x^* if $0 \in ri\partial\mathcal{P}_\mu(x^*)$.

Definition 5.3 (*Strict Second-Order Condition*) We say that the strict second-order condition (*SSOC*) holds at x^* if (T') and (SC) are satisfied at x^* and the Hessian of $L_{\mathcal{U},\mathcal{P}_\mu}(x^*, 0; \cdot)$ is positive definite at $u = 0$.

The computation of $\hat{x}$ in Algorithm 5.1 is not an easy thing. To tackle this difficulty, we proceed in two steps. First we present two propositions, whose ideas come from our latest work in the reference (Huang et al. 2014). For completeness, we shall merely present the results below but omit their details.

Proposition 5.3 *Let $\hat{x} \in \mathcal{W}_r \cup B(x^*, \rho)$. Then $g(\hat{x})$ of (5.20) satisfies*

1. $$g(\hat{x}) = c + \mu D\mathcal{H}(\hat{x})^\star \cdot \left(Q_1(\mathcal{H}(\hat{x}))ZQ_1(\mathcal{H}(\hat{x}))^\top\right),$$

   where Z is a solution of

$$\min_{Z \in S_r^+, \mathrm{tr}Z=1} \left\| c + \mu D\mathcal{H}(\hat{x})^\star \cdot \left(Q_1\left(\mathcal{H}(\hat{x})\right)ZQ_1\left(\mathcal{H}(\hat{x})\right)^\top\right)\right\|^2. \tag{5.22}$$

2. *If (T′) holds at x^*, then the following minimization programming*

$$\min_{Z \in S_r, \mathrm{tr}Z=1} \left\| c + \mu D\mathcal{H}(\hat{x})^\star \cdot \left(Q_1\left(\mathcal{H}(\hat{x})\right)ZQ_1\left(\mathcal{H}(\hat{x})\right)^\top\right)\right\|^2 \tag{5.23}$$

has a unique solution $Z(\hat{x})$ for ρ small enough; moreover the mapping

$$\mathcal{W}_r \cup B(x^*, \rho) \ni \hat{x} \mapsto Z(\hat{x}) \in \mathcal{S}_r$$

is C^∞.

3. *If (T′) and (SC) hold, then $Z(\hat{x}) \succ 0$ for ρ small enough. Consequently $Z(\hat{x})$ is also the unique solution of (5.22) and*

$$g(\hat{x}) \in \mathrm{ri}\partial\mathcal{P}_\mu(\hat{x}). \tag{5.24}$$

Proposition 5.4 *Assume (SSOC) holds at x^*. Then, for ρ small enough, the Dual-step, the $\mathcal{U}$-step and the Update of Algorithm 5.1 are equivalent to the following steps.*
Dual-step. *Compute a solution $Z \in S_r$ of (5.23) and set*

$$G(\hat{x}) := Q_1(\mathcal{H}(\hat{x}))ZQ_1(\mathcal{H}(\hat{x}))^\top.$$

$\mathcal{U}$-step. *Compute a solution $d \in R^m$ of*

$$\begin{aligned} \min_d\ & \langle x - \hat{x} + d, c\rangle + D\mathcal{H}(\hat{x})(x - \hat{x} + d) \cdot G(\hat{x}) \\ & + \tfrac{1}{2} D\mathcal{H}(\hat{x})(x - \hat{x} + d) \cdot H(\mathcal{H}(\hat{x}), G(\hat{x})) \cdot D\mathcal{H}(\hat{x})(x - \hat{x} + d), \\ \text{s.t.}\ & D\mathcal{H}(\hat{x})(x - \hat{x} + d) \in \mathcal{U}_{\lambda_1}(\mathcal{H}(\hat{x})), \end{aligned}$$

Update. *Set $x_+ = x + d$.*

We consider the approximation of $\mathcal{H}(\hat{x})$ by setting the first r eigenvalues equal to $\lambda_1(\mathcal{H}(x))$ without affecting the eigenvectors. For $x \in B(x^*, \rho)$ and ρ small enough, let $Q_{tot}(\mathcal{H}(x))$ be an orthonormal basis of $E_{tot}(\mathcal{H}(x))$ and $\Lambda_{tot}(\mathcal{H}(x)) = Q_{tot}(\mathcal{H}(x))^\top \mathcal{H}(x) Q_{tot}(\mathcal{H}(x))$; then we take as an approximation of $\mathcal{H}(\hat{x})$, the following matrix $\hat{\mathcal{H}}(x) \in \mathcal{M}_r$:

$$\begin{aligned} \hat{\mathcal{H}}(x) = \lambda_1(\mathcal{H}(x)) Q_{tot}(\mathcal{H}(x)) Q_{tot}(\mathcal{H}(x))^\top + \mathcal{H}(x) \\ - Q_{tot}(\mathcal{H}(x)) \Lambda_{tot}(\mathcal{H}(x)) Q_{tot}(\mathcal{H}(x))^\top. \end{aligned}$$

This approximation satisfies

$$\mathcal{H}(x) - \hat{\mathcal{H}}(x) \in \mathrm{span}\partial\lambda_1(\hat{\mathcal{H}}(x)) \subset \ker H(\hat{\mathcal{H}}(x), \hat{G}(x)).$$

Then, Algorithm 5.1 is replaced in the following form.

Algorithm 5.2 Let $x \in B(x^*, \rho)$.
$\mathcal{V}$-Step. Compute $Q_{tot}(\mathcal{H}(x))$, $\Lambda_{tot}(\mathcal{H}(x))$, $\hat{\mathcal{H}}(x)$ and

$$\hat{V}(x) := \mathcal{H}(x) - \hat{\mathcal{H}}(x).$$

Dual-Step. Compute a solution $Z \in S_r$ of

$$\min_{Z \in \mathcal{S}_r, \mathrm{tr} Z = 1} \|c + D\mathcal{H}(\hat{x})^\star \cdot (Q_1(\mathcal{H}(\hat{x}))Z Q_1(\mathcal{H}(\hat{x}))^\top)\|^2$$

and set

$$\hat{G}(x) := Q_1(\hat{\mathcal{H}}(x)) Z Q_1(\hat{\mathcal{H}}(x))^\top.$$

$\mathcal{U}$-Step. Compute the solution $d \in R^m$ of

$$\begin{aligned}
&\min_d \langle \mathcal{H}^{-1}(\hat{V}(x)) + d, c\rangle + \langle D\hat{\mathcal{H}}(x)(\mathcal{H}^{-1}(\hat{V}(x)) + d), \hat{G}(x)\rangle \\
&\quad + \tfrac{1}{2}\langle D\hat{\mathcal{H}}(x)(\mathcal{H}^{-1}(\hat{V}(x)) + d), H(\hat{\mathcal{H}}(x), \hat{G}(x)) \cdot D\hat{\mathcal{H}}(x)(\mathcal{H}^{-1}(\hat{V}(x)) + d)\rangle, \\
&\mathcal{H}(\mathcal{H}^{-1}(\hat{V}(x)) + d) \in \mathcal{U}_{\lambda_1}(\hat{\mathcal{H}}(x)).
\end{aligned} \tag{5.25}$$

Update. Set $x_+ = x + d$.

Remark 5.3 The $\mathcal{U}$-Newton methods discussed here have a key difference with the SQP in Shapiro and Fan (1995). The constrained Newton methods are effective only on the smooth manifold $\mathcal{W}_r$, while SQP is valid over the whole space R^m. In $\mathcal{U}$-Newton methods x is updated to $\hat{x} + u$ with u solving (5.5), where an explicit attempt is made to restore x to the manifold $\mathcal{W}_r$. SQP is intended to achieve both feasibility and optimality asymptotically.

Just as the $\mathcal{U}$-Newton method depends on a choice of $g \in \partial\lambda_1 \circ A(x)$ which defines the $\mathcal{U}$-Lagrangian, SQP depends on the choice of the approximate Lagrange multipliers. Moreover, this choice means that the Newton and SQP directions are the same.

The $\mathcal{U}$-Hessian matrix defined by g reflects the curvature of λ_1 on $\mathcal{W}_r$ near the minimizer x^*, and leads to the following quadratic convergence.

Theorem 5.7 *If (SSOC) holds at x^* then there exists $\rho > 0$ and $C > 0$ such that, for all $x \in B(x^*, \rho)$, x_+ defined by Algorithm 5.2 satisfies:*

$$\|x_+ - x^*\| \le C\|x - x^*\|^2.$$

Proof We adopt the idea in Shapiro and Fan (1995). First we minimize a smooth function $f(x)$ over a smooth manifold $V \subset R^m$. Suppose there exists a neighborhood of x^* such that V can be expressed by a system of smooth equations $g_i(x) = 0, i = 1, \ldots, t$. We adopt the standard Newton method. Let x^k be a current iterate point and β^k the corresponding Lagrange multipliers vector. Then the next iteration is $x^{k+1} = x^k + d^{k+1}$, where d^{k+1} is the solution of the following quadratic programming problem

$$\begin{cases} \min\ d^\top \nabla f(x^k) + \frac{1}{2} d^\top H^k d \\ \text{s.t.}\ \ g_i(x^k) + d^\top \nabla g_i(x^k) = 0, \quad i = 1, \ldots, t, \end{cases} \tag{5.26}$$

and $H^k = \nabla^2_{xx} L(x^k, \beta^k)$ is the Hessian matrix of the Lagrangian

$$L(x, \beta) = f(x) + \sum_{i=1}^{t} \beta_i g_i(x).$$

It is easy to verify that x^{k+1} and the corresponding Lagrange multipliers vector β^{k+1} can be obtained as a solution of the linear equations

$$F(z^k) + \nabla F(z^k)(z - z^k) = 0,$$

where $F(z) = (\nabla_x L(x, \beta), g(x))$, $g(x) = (g_1(x), \ldots, g_t(x))$ and $z = (x, \beta)$.

If the algorithm starts at a point sufficiently close to the optimal solution x^* and the second-order sufficient optimality condition holds at x^*, then the algorithm converges quadratically. In our Algorithm 5.2, the $\mathcal{U}$-Hessian matrix $H(\hat{\mathcal{H}}(x), \hat{G}(x))$ coincides with the above H^k. So the $\mathcal{U}$-step (5.25) can be transformed to the formula (5.26). Under the condition of (SSOC), the quadratic rate of the convergence is obtained. $\square$

6 Conclusions

In this paper, we mainly study the $\mathcal{VU}$ theory about a special class of eigenvalue function, the largest eigenvalues with the matrix convex mapping. $\mathcal{U}$-Lagrangian theory is applied to the class of the functions. when the regular condition holds, we can obtain the first- and second-order derivatives of $\mathcal{U}$-Lagrangian. The second-order analysis of λ_1 using the $\mathcal{U}$-Lagrangian theory is presented. Through the smooth track $\mathcal{X}(u)$ satisfying the transversality condition, there is a second-order expansion of λ_1 along the trajectory $\mathcal{X}(u)$. Furthermore, we present a conceptual algorithm that is proved to have local superlinear convergence. Also, These theoretical finds may be used to deal with some applications: bilinear matrix inequality problems and the maximum eigenvalue with respect to the matrix variables.

The work done on $\mathcal{U}$-Lagrangian applied to the eigenvalue problems is by no means complete. For further work the need can be anticipated: we only give the conceptual algorithm to solve this special class of eigenvalue optimization, next will continue to develop its rapidly convergent executable algorithm and consider how to use the bundle techniques to approximate the proximal points and other $\mathcal{VU}$-related quantities along this line. Moreover, we will strive to extend the $\mathcal{VU}$ theory of convex eigenvalues to nonconvex cases, where its related theory will be researched in later papers.

Acknowledgements The authors would thank the associate editor and the anonymous referees for their valuable and helpful comments and good advice for improving the presentation of this paper. This article is supported by the National Natural Science Foundation of China under Grant Nos. 11301347, 11626053 and 11601389, the Project funded by China Postdoctoral Science Foundation Under No. 2016M601296 and the Fundamental Research Funds for the Central Universities under Project No. 3132016108, the Scientific Research Foundation Funds of DLMU under Project No. 02501102, and Doctoral Foundation of Tianjin Normal University under Project No. 52XB1513.

References

Alizadeh F, Haeberly J-PA, Overton ML (1998) Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. SIAM J Optim 8:746–768

Abraham R, Marsden JE, Ratiu T (1988) Manifolds, tensor analysis, and applications, applied mathematical sciences, vol 75. Springer, Berlin

Apkarian P, Noll D, Thevenet J-B, Tuan HD (2004) A spectral quadratic-SDP method with applications to fixed-order H_2 and H_∞ synthesis. Eur J Control 10:527–538

Bellman R, Fan K (1963) On systems of linear inequalities in Hermitian matrix variables. In: Klee VL (ed.) Convexity volume 7 of proceedings of symposia in pure mathematics, American Mathematical Society, pp 1-11

Bertsekas DP, Nedic A, Ozdaglar AE (2003) Convex analysis and optimization. Athena Scientific and Tsinghua University Press, Belmont

Bonnans JF, Shapiro A (2000) Perturbation analysis of optimization problems. Springer, New York

Cullum J, Donath WE, Wolfe P (1975) The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices. In: Nondifferentiable optimization, mathmatical programming study, vol. 3, pp 35–55

Cox S, Lipton R (1996) Extremal eigenvalue problems for two-phase conductors. Arch Rational Mech Anal 136:101–117

Díaz AR, Kikuchi N (1992) Solution to shape and topology eigenvalue optimization problems using a homogenization method. Int J Numer Methods Eng 35:1487–1502

Fukuda M, Kojima M (2001) Branch-and-cut algorithms for the bilinear matrix inequality eigenvalue problem. Comput Optim Appl 19:79–105

Fletcher R (1985) Semi-definite matrix constraints in optimization. SIAM J Control Optim 23:493–513

Goemans MX (1997) Semidefinite programming in combinatorial optimization. Math Program 79:143–161

Grötschel M, Lovász L, Schrijver A (1983) Geometric algorithms and combinatorial optimization. Springer, Berlin

Helmberg C, Kiwiel KC (2002) A spectral bundle method with bounds. Math Program 93:173–194

Helmberg C, Rendl F (2000) A spectral bundle method for semidefinite programming. SIAM J Optim 10:673–696

Helmberg C, Rendl F, Weismantel R (2000) A semidefinite programming approach to the quadratic knapsack problem. J Comb Optim 4:197–215

Huang M, Pang LP, Xia ZQ (2014) The space decomposition theory for a class of eigenvalue optimizations. Comput Optim Appl 58:423–454

Hiriart-Urruty J-B, Lemaréchal C (1993) Convex analysis and minimization algorithms I-II. In: Grundlehren der mathematischen Wissenschaften, vol. 305–306. Springer-Verlag, Berlin

Hiriart-Urruty J-B, Ye D (1995) Sensitivity analysis of all eigenvalues of a symmetric matrix. Numer Math 70:45–72

Kojima M, Shida M, Shindoh S (1998) Local convergence of predictor-corrector infeasible-interior point algorithms for sdps and sdlcps. Math Program 80:129–161

Lemaréchal C, Oustry F, Sagastizábal C (2000) The $\mathcal{U}$-Lagrangian of a convex function. Trans Am. Math. Soc. 352:711–729

Lewis AS, Overton ML (1996) Eigenvalue optimization. Acta Numer 5:149–190

Lewis AS (1996) Derivatives of spectral functions. Math Oper Res 21:576–588

Lewis AS (2003) Active sets, nonsmoothness and sensitivity. SIAM J Optim 13:702–725

Mangasarian OL (1985) Sufficiency of exact penalty minimization. SIAM J Control Optim 23:30–37

Mifflin R, Sagastizábal C (2005) A $\mathcal{VU}$-algorithm for convex minimization. Math Program Ser B 104:583–608

Noll D, Apkarian P (2005) Spectral bundle method for nonconvex maximum eigenvalue functions: second-order methods. Math Program Ser B 104:729–747

Nesterov Y (1997) Interior-point methods: an old and new approach to nonlinear programming. Math Program 79:285–297

Nesterov Y, Nemirovskii A (1988) A general approach to polynomial-time algorithms design for convex programming. Technical report. Centr. Econ. & Math. Inst., USSR Academy of Sciences, Moscow, USSR

Noll D, Torki M, Apkarian P (2004) Partially augmented Lagrangian method for matrix inequality constraints. SIAM J Optim 15:161–184

Oustry F (1999) The $\mathcal{U}$-Lagrangian of the maximum eigenvalue function. SIAM J Optim 9:526–549
Oustry F (2000) A second-order bundle method to minimize the maximum eigenvalue function. Math Program Ser A 89:1–33
Overton ML (1992) Large-scale optimization of eigenvalues. SIAM J Optim 2:88–120
Overton ML, Womersley RS (1993) Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices. Math Program 62:321–357
Overton ML, Womersley RS (1995) Second derivatives for optimizing eigenvalues of symmetric matrices. SIAM J Matrix Anal Appl 16:697–718
Overton ML, Ye X (1994) Towards second-order methods for structured nonsmooth optimization. In: Gomez S, Hennart J-P (eds) Advances in optimization and numerical analysis, vol 275., Series mathematics and its applications, Kluwer Academic Publishers, Norwell, MA, pp 97–109
Pataki G (1998) On the rank of extreme matrices in semidefinite programming and the multiplicity of optimal eigenvalues. Math Oper Res 23:339–358
Rockafellar RT (1970) Convex analysis. Princeton University Press, New Jersey
Shapiro A, Fan MKH (1995) On eigenvalue optimization. SIAM J Optim 5:552–569
Shapiro A (1997) First and second order analysis of nonlinear semidefinite programs. Math Program Ser B 77:301–320
Thevenet JB, Noll D, Apkarian P (2006) Nonlinear spectral SDP method for BMI-constrained problems: applications to control design. In: Braz J et al (eds) Informatics in control, automation and robotics. Springer, Berlin, pp 61–72
Vandenberghe L, Boyd S (1996) Semidefinite programming. SIAM Rev 38:49–95
Zhao Z, Braams B, Fukuda M, Overton M, Percus J (2004) The reduced density matrix method for electronic structure calculations and the role of three-index representability conditions. J Chem Phys 120:2095–2104