Found. Comput. Math. (2001) 1:205–225

DOI: 10.1007/s102080010008

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Optimal Stability and Eigenvalue Multiplicity

J. V. Burke,[1] A. S. Lewis,[2] and M. L. Overton[3]

[1]Department of Mathematics
University of Washington
Seattle, WA 98195, USA
burke@math.washington.edu.

[2]Department of Combinatorics and Optimization
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada
aslewis@math.uwaterloo.ca.

[3]Courant Institute of Mathematical Sciences
New York University
New York, NY 10012, USA
overton@cs.nyu.edu.

Abstract. We consider the problem of minimizing over an affine set of square matrices the maximum of the real parts of the eigenvalues. Such problems are prototypical in robust control and stability analysis. Under nondegeneracy conditions, we show that the multiplicities of the active eigenvalues at a critical matrix remain unchanged under small perturbations of the problem. Furthermore, each distinct active eigenvalue corresponds to a single Jordan block. This behavior is crucial for optimality conditions and numerical methods. Our techniques blend nonsmooth optimization and matrix analysis.

Date received: June 17, 2000. Final version received: January 4, 2001. Online publication: March 16, 2001. Communicated by Michael Todd.
AMS classification: Primary: 15A42, 93D09, 90C30; Secondary: 65F15, 49K30.
Key words and phrases: Eigenvalue optimization, Robust control, Stability analysis, Spectral abscissa, Nonsmooth analysis, Jordan form.

1. Introduction

Many classical problems of robustness and stability in control theory aim to move the eigenvalues of a parametrized matrix into some prescribed domain of the complex plane (see [**7**] and [**9**], for instance). A particularly important example ("perhaps the most basic problem in control theory" [**3**]) is *stabilization by static output feedback*: given matrices A, B, and C, is there a matrix K such that the matrix $A + BKC$ is *stable* (i.e., has all its eigenvalues in the left half-plane)? In a 1995 survey [**2**], experts in control systems theory described the characterization of those triples (A, B, C) allowing such a K as a "major open problem." With interval bounds on the entries of K, the problem is known to be NP-hard [**3**].

The static output feedback problem is a special case of the general problem of choosing a linearly parametrized matrix X so that its eigenvalues are as far into the left half-plane as possible. This optimizes the asymptotic decay rate of the corresponding dynamical system $\dot{u} = Xu$ (ignoring transient behavior and the possible effects of forcing terms or nonlinearity). Our aim in this paper is to show that the optimal choice of parameters in such problems typically corresponds to patterns of multiple eigenvalues, generalizing an example in [**4**]. This has crucial consequences for the study of optimality conditions and numerical solution techniques.

We denote by $\mathbf{M^n}$ the Euclidean space of n-by-n complex matrices, with inner product $\langle X, Y\rangle = \mathrm{Re}\,\mathrm{tr}(X^*Y)$. The *spectral abscissa* of a matrix X in $\mathbf{M^n}$ is the largest of the real parts of its eigenvalues, denoted $\alpha(X)$. The spectral abscissa is a continuous function, but it is not smooth, convex, or even locally Lipschitz. We call an eigenvalue λ of X *active* if $\mathrm{Re}\,\lambda = \alpha(X)$, and *nonderogatory* if $X - \lambda I$ has rank $n - 1$ (or, in other words, λ corresponds to a single Jordan block), and we say X *has nonderogatory spectral abscissa* if all its active eigenvalues are nonderogatory. We call X *nonderogatory* if all its eigenvalues are nonderogatory. It is elementary that the set of nonderogatory matrices is dense and open in $\mathbf{M^n}$. In a precise sense, derogatory matrices are "rare": within the manifold of matrices with any given set of eigenvalues and multiplicities, the subset of derogatory matrices has strictly smaller dimension [**1**].

Informally, we consider the *active manifold* $\mathcal{M}$ at a matrix X with nonderogatory spectral abscissa as the set of matrices close to X with active eigenvalues having the same multiplicities as those of X. (These eigenvalues are then necessarily nonderogatory.) Later we give a formal definition of $\mathcal{M}$, and we show that α is smooth on $\mathcal{M}$. We denote the tangent and normal spaces to $\mathcal{M}$ at X by $T_{\mathcal{M}}(X)$ and $N_{\mathcal{M}}(X)$, respectively.

Given a real-valued smooth function defined on some manifold in a Euclidean space, we say a point in the manifold is a *strong minimizer* if, in the manifold, the function grows at least quadratically near the point. If the function is smooth on a neighborhood of the manifold, this is equivalent to the gradient of the function being normal to the manifold at the point and the Hessian of the function being positive definite on the tangent space to the manifold at the point.

Given a Euclidean space $\mathbf{E}$ (a real inner product space) and a linear map $\Phi\colon \mathbf{E} \to \mathbf{M^n}$, we denote the range and nullspace of Φ by $R(\Phi)$ and $N(\Phi)$, respectively. The adjoint of Φ we denote $\Phi^*\colon \mathbf{M^n} \to \mathbf{E}$. The map Φ is *transversal* to the active manifold $\mathcal{M}$ at X if

$$N(\Phi^*) \cap N_{\mathcal{M}}(X) = \{0\}.$$

For a fixed matrix A in $\mathbf{M^n}$ we now consider the problem of minimizing the spectral abscissa over the affine set $A + R(\Phi)$:

$$\inf_{A+R(\Phi)} \alpha. \tag{1.1}$$

For simplicity, we restrict attention to linearly parametrized spectral abscissa minimization problems. It is not hard to extend the approach to smoothly parametrized problems.

Suppose the matrix X is locally optimal for the spectral abscissa minimization problem (1.1). The nonsmooth nature of this problem suggests an approach to optimality conditions via modern variational analysis, for which a comprehensive reference is Rockafellar and Wets' recent work [**10**]: we rely on this reference throughout. If X has nonderogatory spectral abscissa and the map Φ is transversal to the active manifold $\mathcal{M}$ at X, then we shall see that the following first-order necessary optimality condition must hold:

$$0 \in \Phi^* \partial\alpha(X), \tag{1.2}$$

where ∂ denotes the subdifferential (the set of subgradients in the sense of [**10**]). Furthermore, X must be a critical point for the smooth function $\alpha|_{\mathcal{M}\cap(A+R(\Phi))}$, and indeed a local minimizer.

These necessary optimality conditions are easily seen not to be sufficient, and, just as in classical nonlinear programming, to study perturbation theory we need stronger assumptions. We therefore make the following definition:

Definition 1.3. Suppose the map Φ is injective. The matrix X is a **nondegenerate critical point** for the spectral abscissa minimization problem

$$\inf_{A+R(\Phi)} \alpha$$

if it satisfies the following conditions:

(i) X has nonderogatory spectral abscissa;
(ii) Φ is transversal to the active manifold $\mathcal{M}$ at X;
(iii) $0 \in \operatorname{ri} \Phi^* \partial\alpha(X)$;
(iv) X is a strong minimizer of $\alpha|_{\mathcal{M}\cap(A+R(\Phi))}$.

(Here, ri denotes the relative interior of a convex set.) Under these conditions we show that if we perturb the matrix A and the map Φ in the spectral abscissa minimization problem, then the new problem has a nondegenerate critical point close to X with active eigenvalues having the same multiplicities as those of X. In other words, the "active" Jordan structure at a nondegenerate critical point persists under small perturbations to the problem.

2. Arnold Form and the Active Manifold

Henceforth we fix a matrix $\bar{X}$ in $\mathbf{M^n}$ with nonderogatory spectral abscissa. In terms of the Jordan form of $\bar{X}$ this means there is an integer $p > 0$ and block size integers $m_0 \geq 0$, $m_1, m_2, \ldots, m_p > 0$, with sum n, a vector $\bar{\lambda}$ in $\mathbf{C^p}$ whose components (the active eigenvalues) have distinct imaginary parts and real parts all equal to the spectral abscissa $\alpha(\bar{X})$, a matrix $\bar{B}$ in $\mathbf{M^{m_0}}$ (the "inactive part" of $\bar{X}$) with spectral abscissa strictly less than $\alpha(\bar{X})$, and an invertible matrix $\bar{P}$ in $\mathbf{M^n}$ that reduces $\bar{X}$ to the block-diagonal form

$$\bar{P}\bar{X}\bar{P}^{-1} = \mathrm{Diag}(\bar{B}, 0, 0, \ldots, 0) + \sum_{j=1}^{p}(J_{j1} + \bar{\lambda}_j J_{j0}).$$

Here, the expression $\mathrm{Diag}(\cdot, \cdot, \ldots, \cdot)$ denotes a block-diagonal matrix with block sizes $m_0, m_1, \ldots, m_p$, and

$$J_{jq} = \mathrm{Diag}(0, 0, \ldots, 0, J_{m_j}^q, 0, \ldots, 0),$$

where J_m^q denotes the qth power of the elementary m-by-m Jordan block

$$J_m = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

(We make the natural interpretation that J_m^0 is the identity matrix.) We do not assume the inactive part $\bar{B}$ is in Jordan form or that it is nonderogatory.

We mostly think of $\mathbf{M^n}$ as a Euclidean space, but by considering the sesquilinear form $\langle X, Y\rangle_{\mathbf{C}} = \mathrm{tr}(X^*Y)$, we can also consider it as a complex inner product space, which we denote $\mathbf{M^n_C}$. It is easy to check that the matrices J_{jq} are mutually orthogonal in this space:

$$\mathrm{tr}(J_{jq}^* J_{kr}) = \begin{cases} m_j - q & \text{if } j = k \text{ and } q = r, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

We denote $\sqrt{-1}$ by i.

Our key tool is the following result from [**1**]: it describes a smooth reduction of any matrix close to $\bar{X}$ to a certain normal form.

Theorem 2.2 (Arnold Form). *There is a neighborhood* Ω *of* $\bar{X}$ *in* $\mathbf{M^n}$, *and smooth maps* $P\colon \Omega \to \mathbf{M^n}$, $B\colon \Omega \to \mathbf{M^{m_0}}$, *and* $\lambda_{jq}\colon \Omega \to \mathbf{C}$, *such that* $P(\bar{X}) = \bar{P}$, $B(\bar{X}) = \bar{B}$,

$$\lambda_{jq}(\bar{X}) = \begin{cases} \bar{\lambda}_j & \text{if } q = 0, \\ 0 & \text{otherwise}, \end{cases}$$

for $q = 0, 1, \ldots, m_j - 1$ *and* $j = 1, 2, \ldots, p$, *and*

$$P(X)XP(X)^{-1} = \mathrm{Diag}(B(X), 0, 0, \ldots, 0) + \sum_{j=1}^{p} \left(J_{j1} + \sum_{q=0}^{m_j - 1} \lambda_{jq}(X) J_{jq}^{*} \right).$$

In fact the functions λ_{jq} in the Arnold form are uniquely defined near $\bar{X}$ (although the maps P and B are not unique). To see this we again follow [**1**], considering the *orbit* of a matrix Z in $\mathbf{M^m}$ (the set of matrices similar to Z), which we denote orb Z. Two nonderogatory matrices W and Z are similar if and only if they have the same eigenvalues (with the same multiplicities), by considering their Jordan forms. Equivalently, nonderogatory W and Z are similar if and only if their characteristic polynomials coincide or, in other words,

$$p_r(W) = p_r(Z) \qquad (r = 1, 2, \ldots, m),$$

where $p_r\colon \mathbf{M^m_C} \to \mathbf{C}$ is the homogeneous polynomial of degree r (for $r = 0, 1, \ldots, m$) defined by

$$\det(tI - Z) = \sum_{r=0}^{m} p_r(Z) t^{m-r} \qquad (t \in \mathbf{C}). \tag{2.3}$$

(For example, $p_0 \equiv 1$, $p_1 = -\,\mathrm{tr}$ and $p_m = (-1)^m \det$.) Hence if $Z \in \mathbf{M^m}$ is nonderogatory then we can define the orbit of Z locally by

$$\Omega \cap \mathrm{orb}\, Z = \{W \in \Omega\colon\ p_r(W) = p_r(Z)\ (r = 1, 2, \ldots, m)\}$$

for some neighborhood Ω of Z.

In the case $Z = J_m$ we obtain

$$\Omega \cap \mathrm{orb}\, J_m = \{W \in \Omega\colon\ p_r(W) = 0\ (r = 1, 2, \ldots, m)\}. \tag{2.4}$$

Furthermore, by differentiating (2.3) with respect to Z at $Z = J_m$ we obtain, for sufficiently large t,

$$\sum_{r=0}^{m} \nabla p_r(J_m) t^{m-r} = -(tI - J_m)^{-1} \det(tI - J_m) = -\sum_{r=0}^{m-1} t^{m-r-1} J_m^r,$$

so

$$\nabla p_r(J_m) = -J_m^{r-1} \qquad (r = 1, 2, \ldots, m). \tag{2.5}$$

(For a more general version of this result, see [**5**, Lemma 7.1].) These gradients are linearly independent, so (2.4) describes orb J_m locally as a manifold of codimension m, with normal space

$$N_{\mathrm{orb}\, J_m}(J_m) = \mathrm{span}\{(J_m^{r-1})^*\colon\ r = 1, 2, \ldots, m\}$$

(cf. [**1**, Theorem 4.4]).

The following result is an elementary presentation of the basic example of Arnold's key idea.

Lemma 2.6. *Any matrix Z close to J_m is similar to a unique matrix* (*depending smoothly on Z*) *in $J_m + N_{\mathrm{orb}\, J_m}(J_m)$.*

Proof. We need to show the system of equations

$$p_r\left(J_m + \sum_{k=1}^{m} \frac{\lambda_k}{m-k+1}\left(J_m^{k-1}\right)^*\right) = p_r(Z) \qquad (r = 1, 2, \ldots, m)$$

has a unique small solution $\lambda(Z) \in \mathbf{C^m}$ for Z close to J_m, where λ is smooth and $\lambda(J_m) = 0$. But at $Z = J_m$ we have the solution $\lambda = 0$ and the Jacobian of the system is minus the identity, by our gradient calculation (2.5) and a special case of the orthogonality relationship (2.1). The result now follows from the inverse function theorem. □

Theorem 2.7 (Arnold Form Is Well Defined). *The functions λ_{jq} in Theorem* 2.2 *are uniquely defined on a neighborhood of $\bar{X}$.*

Proof. Suppose we have another Arnold form defined, analogously to Theorem 2.2, on a neighborhood $\tilde{\Omega}$ of $\bar{X}$ by smooth maps $\tilde{P}$, $\tilde{B}$, and $\tilde{\lambda}_{jq}$ for each j and q. By continuity of the eigenvalues we know the matrices

$$J_{m_j}^* + \sum_{q=0}^{m_j-1} \lambda_{jq}(X) J_{m_j}^q \qquad \text{and} \qquad J_{m_j}^* + \sum_{q=0}^{m_j-1} \tilde{\lambda}_{jq}(X) J_{m_j}^q$$

have the same eigenvalues for X close to $\bar{X}$ and $j = 1, 2, \ldots, p$. Since they are both nonderogatory they are similar. Hence, by the preceding lemma, $\lambda_{jq}(X) = \tilde{\lambda}_{jq}(X)$ for all j and q, as required. □

Now we can give a precise definition of the active manifold we discussed in the previous section.

Definition 2.8. With the notation of the Arnold form above, the **active manifold** $\mathcal{M}$ at $\bar{X}$ is the set of matrices X in $\Omega \subset \mathbf{M^n}$ satisfying

$$\begin{aligned} \lambda_{jq}(X) &= 0 \qquad (q = 1, 2, \ldots, m_j - 1,\ j = 1, 2, \ldots, p), \\ \mathrm{Re}\,\lambda_{j0}(X) &= \beta \qquad (j = 1, 2, \ldots, p), \end{aligned}$$

for some real β.

The following result describes the active manifold more intuitively:

Theorem 2.9 (Structure of Active Manifold). *A matrix X close to $\bar{X}$ lies in the active manifold $\mathcal{M}$ if and only if there is a matrix P close to $\bar{P}$, a matrix B close to $\bar{B}$ with spectral abscissa strictly less than $\alpha(X)$, and a vector λ close to $\bar{\lambda}$ whose components have distinct imaginary parts and real parts all equal to $\alpha(X)$, such that X has the same "active Jordan structure" as $\bar{X}$:*

$$PXP^{-1} = \mathrm{Diag}(B, 0, 0, \ldots, 0) + \sum_{j=1}^{p}(J_{j1} + \lambda_j J_{j0}).$$

Proof. The "only if" direction follows immediately from the definition. The proof of the converse is analogous to that of Theorem 2.7 (the Arnold form is well defined). □

Crucial to our analysis will be the following observation:

Proposition 2.10. *The spectral abscissa is smooth on the active manifold.*

Proof. This follows immediately from the definition, since

$$\alpha(X) = \mathrm{Re}\,\lambda_{j0}(X) \qquad (j = 1, 2, \ldots, p) \tag{2.11}$$

for all matrices X in the active manifold $\mathcal{M}$. □

Since we are concerned with optimality conditions involving the spectral abscissa restricted to the active manifold, we need to calculate the tangent and normal spaces to $\mathcal{M}$ at $\bar{X}$. In the next result we therefore compute the gradients of the functions λ_{jq} (cf. [**8**, Theorem 2.4]).

Lemma 2.12. *On the complex inner product space $\mathbf{M}^{\mathbf{n}}_{\mathbf{C}}$, the gradient of the function $\lambda_{jq}\colon \Omega \to \mathbf{C}$ at $\bar{X}$ is given by*

$$(\nabla\lambda_{jq}(\bar{X}))^* = (m_j - q)^{-1}\bar{P}^{-1}J_{jq}\bar{P}.$$

Furthermore, on the Euclidean space $\mathbf{M}^{\mathbf{n}}$, for any complex μ, the function $\mathrm{Re}(\mu\lambda_{jq})\colon \Omega \to \mathbf{R}$ *has gradient*

$$(\nabla(\mathrm{Re}(\mu\lambda_{jq}))(\bar{X}))^* = (m_j - q)^{-1}\bar{P}^{-1}\mu J_{jq}\bar{P},$$

for $q = 0, 1, \ldots, m_j - 1$ and $j = 1, 2, \ldots, p$.

Proof. Fix an integer k such that $1 \le k \le p$ and an integer r such that $0 \le r \le m_k - 1$. In [**1**, §4.1], Arnold observes that, in the space $\mathbf{M}^{\mathbf{n}}_{\mathbf{C}}$, the matrix J^*_{kr}

is orthogonal to the orbit of $\bar{X}$ at $\bar{P}\bar{X}\bar{P}^{-1}$. Hence, for matrices X close to $\bar{X}$, we know

$$\begin{aligned}
&\operatorname{tr}(J_{kr}(P(X)XP(X)^{-1} - \bar{P}\bar{X}\bar{P}^{-1}))\\
&\qquad = \operatorname{tr}(J_{kr}(P(X)(X-\bar{X})P(X)^{-1} + P(X)\bar{X}P(X)^{-1} - \bar{P}\bar{X}\bar{P}^{-1}))\\
&\qquad = \operatorname{tr}(J_{kr}P(X)(X-\bar{X})P(X)^{-1}) + o(X-\bar{X})\\
&\qquad = \operatorname{tr}(J_{kr}\bar{P}(X-\bar{X})\bar{P}^{-1}) + o(X-\bar{X}).
\end{aligned}$$

But from the Arnold form (Theorem 2.2) we also know

$$\begin{aligned}
&P(X)XP(X)^{-1} - \bar{P}\bar{X}\bar{P}^{-1}\\
&\qquad = \operatorname{Diag}(B(X)-\bar{B}, 0, 0, \ldots, 0) + \sum_{j=1}^{p}\sum_{q=0}^{m_j-1}(\lambda_{jq}(X) - \lambda_{jq}(\bar{X}))J_{jq}^*.
\end{aligned}$$

Taking the inner product with J_{kr} and using the previous expression and the orthogonality relation (2.1) shows

$$(m_k - r)(\lambda_{kr}(X) - \lambda_{kr}(\bar{X})) = \operatorname{tr}(J_{kr}\bar{P}(X-\bar{X})\bar{P}^{-1}) + o(X-\bar{X}),$$

which proves the required expression for $\nabla\lambda_{jq}$. The second expression follows after multiplication by μ and taking real parts. □

We deduce the following result:

Theorem 2.13 (Tangent Space). *The tangent space $T_{\mathcal{M}}(\bar{X})$ to the active manifold at $\bar{X}$ is the set of matrices Z in $\mathbf{M^n}$ satisfying*

$$\begin{aligned}
\operatorname{tr}(\bar{P}^{-1}J_{jq}\bar{P}Z) &= 0 \qquad (q = 1, 2, \ldots, m_j - 1,\ j = 1, 2, \ldots, p),\\
m_j^{-1}\operatorname{Re}\operatorname{tr}(\bar{P}^{-1}J_{j0}\bar{P}Z) &= \beta \qquad (j = 1, 2, \ldots, p),
\end{aligned}$$

for some real β.

Proof. By definition, the active manifold is the set of matrices X in $\Omega \subset \mathbf{M^n}$ satisfying

$$\begin{aligned}
\operatorname{Re}\lambda_{jq}(X) = \operatorname{Re}(i\lambda_{jq}(X)) &= 0 \qquad (q = 1, 2, \ldots, m_j - 1,\ j = 1, 2, \ldots, p),\\
\operatorname{Re}\lambda_{j0}(X) - \operatorname{Re}\lambda_{10}(X) &= 0 \qquad (j = 2, 3, \ldots, p).
\end{aligned}$$

By Lemma 2.12 the gradients of these constraint functions are linearly independent at $X = \bar{X}$, so the tangent space consists of those Z satisfying

$$\begin{aligned}
\operatorname{Re}\operatorname{tr}(\bar{P}^{-1}J_{jq}\bar{P}Z) = \operatorname{Re}\operatorname{tr}(\bar{P}^{-1}iJ_{jq}\bar{P}Z) &= 0\\
&\qquad (q = 1, 2, \ldots, m_j - 1,\ j = 1, 2, \ldots, p),\\
\operatorname{Re}\operatorname{tr}(m_j^{-1}\bar{P}^{-1}J_{j0}\bar{P}Z) - \operatorname{Re}\operatorname{tr}(m_1^{-1}\bar{P}^{-1}J_{10}\bar{P}Z) &= 0 \qquad (j = 2, 3, \ldots, p),
\end{aligned}$$

as claimed. □

To simplify our expressions for the corresponding normal space and related sets of subgradients, we introduce three polyhedral sets in the space

$$\boldsymbol{\Theta} = \{\theta = (\theta_{jq}) \colon \theta_{j0} \in \mathbf{R},\ \theta_{jq} \in \mathbf{C}\ (q = 1, 2, \ldots, m_j - 1,\ \ j = 1, 2, \ldots, p)\}.$$

Specifically, we define

$$\begin{aligned}
\Theta_0 &= \left\{\theta \in \boldsymbol{\Theta} \colon \sum_{j=1}^{p} m_j \theta_{j0} = 0\right\}, \\
\Theta_1 &= \left\{\theta \in \boldsymbol{\Theta} \colon \sum_{j=1}^{p} m_j \theta_{j0} = 1\right\}, \\
\Theta_1^{\geq} &= \{\theta \in \Theta_1 \colon \theta_{j0} \geq 0,\ \operatorname{Re}\theta_{j1} \geq 0\ (j = 1, 2, \ldots, p)\}.
\end{aligned}$$

Notice that the set Θ_1 is the affine span of the set $\Theta_1^{\geq}$, and is a translate of the subspace Θ_0. We also define the injective linear map

$$\Lambda \colon \boldsymbol{\Theta} \to \mathbf{M^n}$$

by

$$\Lambda\theta = \sum_{j=1}^{p} \sum_{q=0}^{m_j - 1} \theta_{jq} J_{jq}^*.$$

Corollary 2.14 (Normal Space). *In the Euclidean space* $\mathbf{M^n}$, *the normal space to the active manifold* $\mathcal{M}$ *at the matrix* $\bar{X}$ *is given by*

$$N_{\mathcal{M}}(\bar{X}) = \bar{P}^*(\Lambda\Theta_0)\bar{P}^{-*}.$$

Proof. The proof of the previous result shows that in $\mathbf{M^n}$, the tangent space $T_{\mathcal{M}}(\bar{X})$ is the orthogonal complement of the set of matrices

$$\begin{aligned}
&\bar{P}^*\{J_{jq}^*, -iJ_{jq}^* \colon q = 1, 2, \ldots, m_j - 1,\ \ j = 1, 2, \ldots, p\}\bar{P}^{-*} \\
&\quad \cup\ \bar{P}^*\{m_j^{-1} J_{j0}^* - m_1^{-1} J_{10}^* \colon 2 \leq j \leq p\}\bar{P}^{-*}.
\end{aligned}$$

Using the elementary relationship, for arbitrary vectors $a_1, a_2, \ldots, a_p$:

$$\operatorname{span}\{a_j - a_1 \colon 2 \leq j \leq p\} = \left\{\sum_{j=1}^{p} \mu_j a_j \colon \sum_{j=1}^{p} \mu_j = 0\right\},$$

we see that the span of the second set is just

$$\bar{P}^* \left\{\sum_{j=1}^{p} \theta_{j0} J_{j0} \colon \sum_{j=1}^{p} m_j \theta_{j0} = 0\right\} \bar{P}^{-*},$$

and the result now follows. □

It is highly instructive to compare this expression for the normal space to the active manifold at the matrix $\bar{X}$ to the subdifferential (the set of subgradients) of the spectral abscissa at the same matrix. In [**5**] Burke and Overton obtain the following result:

Theorem 2.15 (Spectral Abscissa Subgradients). *The spectral abscissa α is regular at $\bar{X}$, in the sense of* [**10**], *and its subdifferential is*

$$\partial\alpha(\bar{X}) = \bar{P}^*(\Lambda\Theta_1^{\geq})\bar{P}^{-*}.$$

Since the set $\Lambda\Theta_1^{\geq}$ is a translated polyhedral cone with affine span a translate of $\Lambda\Theta_0$, we deduce the following consequence:

Corollary 2.16. *The subdifferential $\partial\alpha(\bar{X})$ is a translated polyhedral cone with affine span a translate of the normal space $N_{\mathcal{M}}(\bar{X})$.*

Notice the following particular elements of the subdifferential:

Corollary 2.17. *For every index $1 \leq j \leq p$ we have*

$$\nabla(\operatorname{Re}\lambda_{j0})(\bar{X}) \in \partial\alpha(\bar{X}).$$

Proof. This follows from Lemma 2.12 and Theorem 2.15 (spectral abscissa subgradients). □

3. Linearly Parametrized Problems

We are interested in optimality conditions for local minimizers of the spectral abscissa over a given affine set of matrices. We can write this affine set as a translate of the range of a linear map $\bar{\Phi}\colon \mathbf{E} \to \mathbf{M^n}$, where $\mathbf{E}$ is a Euclidean space, so our problem becomes

$$\inf_{\bar{A}+R(\bar{\Phi})} \alpha,$$

where $\bar{A}$ is a fixed matrix in $\mathbf{M^n}$. Equivalently, if we define a function $\bar{f}\colon \mathbf{E} \to \mathbf{R}$ by

$$\bar{f}(z) = \alpha(\bar{A} + \bar{\Phi}(z)),$$

then we are interested in local minimizers of $\bar{f}$. Suppose therefore that the point $\bar{z}$ in $\mathbf{E}$ satisfies

$$\bar{A} + \bar{\Phi}(\bar{z}) = \bar{X}.$$

We henceforth make the following regularity assumption:

Assumption 3.1. *The matrix $\bar{X}$ has nonderogatory spectral abscissa and the map $\bar{\Phi}$ is injective and transversal to the active manifold $\mathcal{M}$ at $\bar{X}$.*

Under this assumption we can compute the normal space to the *parameter-active manifold* $\bar{\Phi}^{-1}(\mathcal{M} - \bar{A})$ (the set of those parameters z for which $\bar{A} + \bar{\Phi}(z)$ lies in the active manifold), and we can apply a nonsmooth chain rule to compute the subdifferential of $\bar{f}$.

Proposition 3.2 (Necessary Optimality Condition).

$$N_{\bar{\Phi}^{-1}(\mathcal{M}-\bar{A})}(\bar{z}) = \bar{\Phi}^* N_{\mathcal{M}}(\bar{X}),$$

and

$$\partial \bar{f}(\bar{z}) = \bar{\Phi}^* \partial\alpha(\bar{X}).$$

Furthermore, if $\bar{X}$ is a local minimizer of α on $\bar{A} + R(\bar{\Phi})$, then the first-order optimality condition

$$0 \in \bar{\Phi}^* \partial\alpha(\bar{X})$$

must hold.

Proof. The first equation is a classical consequence of transversality (or see, e.g., [**10**, Theorem 6.14]). Now, note that α is regular at $\bar{X}$, by Theorem 2.15 (spectral abscissa subgradients). Hence the horizon subdifferential $\partial^\infty\alpha(\bar{X})$ is just the recession cone of the subdifferential $\partial\alpha(X)$, and is therefore a subset of the normal space $N_{\mathcal{M}}(\bar{X})$, by Corollary 2.14 (normal space). Thus transversality implies the condition

$$N(\bar{\Phi}^*) \cap \partial^\infty\alpha(\bar{X}) = \{0\},$$

so we can apply the chain rule [**10**, Theorem 10.6] to deduce the second equation. If $\bar{X}$ is a local minimizer, then $\bar{z}$ is a local minimizer of $\bar{f}$, so satisfies $0 \in \partial \bar{f}(\bar{z})$, whence the optimality condition. □

We know, by Proposition 2.10, that the restriction of the spectral abscissa to the set of feasible matrices in the active manifold is a smooth function. Clearly, if $\bar{X}$ is a local minimizer for our problem, then in particular it must be a critical point of this smooth function. The next result states that, conversely, such critical points must satisfy a weak form of the necessary optimality condition.

Theorem 3.3 (Restricted Optimality Condition). *The matrix $\bar{X}$ is a critical point of the smooth function $\alpha|_{\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))}$ if and only if it satisfies the condition*

$$0 \in \operatorname{aff} \bar{\Phi}^* \partial\alpha(\bar{X}).$$

Proof. First note $\bar{X}$ is a critical point if and only if $\bar{z}$ is a critical point of $\bar{f}|_{\bar{\Phi}^{-1}(\mathcal{M}-\bar{A})}$. Pick any index $1 \le j \le p$, and define a smooth function coinciding with $\bar{f}$ on the parameter-active manifold:

$$g(z) = \operatorname{Re}\lambda_{j0}(\bar{A} + \bar{\Phi}(z)) \quad \text{for } z \text{ close to } \bar{z}.$$

By the chain rule we know $\nabla g(\bar{z}) = \bar{\Phi}^*(Y)$, where

$$Y = \nabla(\operatorname{Re}\lambda_{j0})(\bar{X}) \in \partial\alpha(\bar{X}),$$

by Corollary 2.17. Clearly $\bar{z}$ is a critical point if and only if this gradient lies in the normal space to the parameter-active manifold:

$$\nabla g(\bar{z}) \in N_{\bar{\Phi}^{-1}(\mathcal{M}-\bar{A})}(\bar{z}).$$

By Proposition 3.2 (necessary optimality condition), this is equivalent to

$$\bar{\Phi}^*(Y) \in \bar{\Phi}^*(N_{\mathcal{M}}(\bar{X})),$$

or, in other words,

$$0 \in \bar{\Phi}^*(Y + N_{\mathcal{M}}(\bar{X})).$$

But, since $Y \in \partial\alpha(\bar{X})$, by Corollary 2.16 we know

$$\operatorname{aff}\partial\alpha(\bar{X}) = Y + N_{\mathcal{M}}(\bar{X}),$$

so the result follows. □

We call critical points of $\alpha|_{\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))}$ *active-critical*.

The matrix Y appearing in the above proof is a "dual matrix," analogous to the construction in [**5**, §10]. It is illuminating to consider an alternative proof, from a slightly different perspective, where Y appears as a Lagrange multiplier. By Theorem 2.15 (spectral abscissa subgradients), the condition $0 \in \operatorname{aff}\bar{\Phi}^*\partial\alpha(\bar{X})$ holds if and only if there exists $\bar{\theta} \in \mathbf{\Theta}$ satisfying

$$\sum_{j=1}^{p} m_j\bar{\theta}_{j0} = 1, \tag{3.4}$$

and a matrix $\bar{Y}$ in $\mathbf{M^n}$ satisfying

$$\bar{\Phi}^*(\bar{Y}) = 0 \tag{3.5}$$

and

$$\bar{Y} = \bar{P}^*\Lambda(\bar{\theta})\bar{P}^{-*}. \tag{3.6}$$

On the other hand, by the definition of the active manifold (Definition 2.8), the matrix $\bar{X}$ is a critical point of the function $\alpha|_{\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))}$ if and only if the smooth equality-constrained minimization problem

$$SAM(\bar{A},\bar{\Phi})\quad\begin{cases}\text{inf} & \beta\\ \text{subject to} & m_j(\operatorname{Re}\lambda_{j0}(X)-\beta) = 0,\\ & (m_j-q)\lambda_{jq}(X) = 0,\\ & \quad (q=1,2,\dots,m_j-1,\\ & \quad\; j=1,2,\dots,p)\\ & X-\bar{\Phi}(z) = \bar{A},\\ & \beta\in\mathbf{R},\ X\in\mathbf{M^n},\ z\in\mathbf{E},\end{cases}$$

has a critical point at $(\beta, X, z) = (\alpha(\bar{X}), \bar{X}, \bar{z})$. Consider the Lagrangian

$$\begin{aligned}L(\beta, X, z, \theta, Y)\\ &= \beta + \operatorname{Re}\operatorname{tr}(Y^*(X-\bar{\Phi}(z))\\ &\quad+\sum_{j=1}^{p}\left(\theta_{j0}m_j(\operatorname{Re}\lambda_{j0}(X)-\beta)+\sum_{q=1}^{m_j-1}\operatorname{Re}(\theta_{jq}^*(m_j-q)\lambda_{jq}(X))\right).\end{aligned}$$

We can write an arbitrary linear combination of the constraints as

$$L(\beta, X, z, \theta, Y) - \beta.$$

Differentiating this linear combination at the critical point with respect to β, z, and X (using Lemma 2.12) shows

$$\sum_{j=1}^{p} m_j\theta_{j0} = 0, \qquad \bar{\Phi}^*(Y) = 0, \qquad \text{and} \qquad Y = \bar{P}^*\Lambda(\theta)\bar{P}^{-*}, \tag{3.7}$$

respectively. This implies $Y = 0$ and hence $\theta = 0$, by transversality and Corollary 2.14 (normal space), so the constraints have linearly independent gradients at this point.

Thus the point of interest is critical if and only if there exist Lagrange multipliers $\bar{\theta}$ and $\bar{Y}$ (necessarily unique) such that the Lagrangian $L(\beta, X, z, \bar{\theta}, \bar{Y})$ is critical. Differentiating L with respect to β, z, and X again gives conditions (3.4), (3.5), and (3.6), respectively, so the result follows.

The Lagrange multipliers $\bar{\theta}$ and $\bar{Y}$ also allow us to check the second-order conditions for the above optimization problem $SAM(\bar{A}, \bar{\Phi})$. Summarizing this discussion, we have the following result:

Theorem 3.8 (Active-Critical Points). *The following conditions are all equivalent*:

(i) *$\bar{X}$ is critical for the function $\alpha|_{\mathcal{M}\cap\bar{A}+R(\bar{\Phi}))}$.*
(ii) *$(\alpha(\bar{X}), \bar{X}, \bar{z})$ is critical for the problem $SAM(\bar{A}, \bar{\Phi})$.*

(iii) *Equations* (3.4), (3.5), *and* (3.6) *have a solution* $(\bar{\theta}, \bar{Y})$.
(iv) $0 \in \operatorname{aff} \bar{\Phi}^* \partial\alpha(\bar{X})$.

Suppose these conditions hold. Then the Lagrange multiplier $(\bar{\theta}, \bar{Y})$ *is unique, and we have*

$$0 \in \bar{\Phi}^* \partial\alpha(\bar{X}) \Leftrightarrow \bar{\theta} \in \Theta_1^{\geq}$$

and

$$0 \in \operatorname{ri} \bar{\Phi}^* \partial\alpha(\bar{X}) \Leftrightarrow \bar{\theta} \in \operatorname{ri} \Theta_1^{\geq}.$$

Furthermore, $\bar{X}$ *is a strong minimizer of* $\alpha|_{\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))}$ *if and only if the Hessian of the Lagrangian, namely*

$$\nabla^2 \sum_{j=1}^{p} \sum_{q=0}^{m_j-1} \operatorname{Re}(\bar{\theta}_{jq}^*(m_j - q)\lambda_{jq}(X)),$$

is positive definite on the space $T_{\mathcal{M}}(\bar{X}) \cap R(\bar{\Phi})$.

Proof. The first four equivalences and the uniqueness of the Lagrange multiplier follow from the previous discussion. The next two equivalences are a consequence of Theorem 2.15 (spectral abscissa subgradients) and the uniqueness of the Lagrange multiplier.

Only the second-order condition remains to be proved. Clearly $\bar{X}$ is a strong minimizer of $\alpha|_{\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))}$ if and only if $(\alpha(\bar{X}), \bar{X}, \bar{z})$ is a strong minimizer of β restricted to the feasible region of the problem $SAM(\bar{A}, \bar{\Phi})$. By standard second-order optimality theory this is equivalent to the Lagrangian having positive-definite Hessian on the tangent space to the feasible region which, by transversality, is just $T_{\mathcal{M}}(\bar{X}) \cap R(\bar{\Phi})$. □

Recall that the tangent space to the active manifold $T_{\mathcal{M}}(\bar{X})$ is given by Theorem 2.13. Thus, in principle, the second-order optimality condition above is checkable, since the second derivatives at $\bar{X}$ of the functions λ_{jq} in the Arnold form (Theorem 2.2) are computable (see [**8**], for example): we do not pursue this here.

4. Perturbation

We are now ready to prove our main result. It uses the notion of a nondegenerate critical point (Definition 1.3).

Theorem 4.1 (Persistence of Active Jordan Form). *Suppose the matrix* $\bar{X}$ *is a nondegenerate critical point for the spectral abscissa minimization problem*

$$\inf_{\bar{A}+R(\bar{\Phi})} \alpha,$$

where the linear map $\bar{\Phi}$ is injective. Then, for matrices A close to $\bar{A}$ and maps Φ close to $\bar{\Phi}$, there is a unique critical point of the smooth function $\alpha|_{\mathcal{M}\cap(A+R(\Phi))}$ close to $\bar{X}$. This point is a nondegenerate critical point for the perturbed problem

$$\inf_{A+R(\Phi)} \alpha,$$

depending smoothly on the data A and Φ: in particular, it has the same active Jordan structure as $\bar{X}$.

Proof. The matrix $\bar{X}$ satisfies the first- and second-order conditions of Theorem 3.8 (active-critical points), so using the notation of that result, we can apply standard nonlinear programming sensitivity theory (see, e.g., [**6**, §2.4]) to deduce the existence of a smooth function

$$(A, \Phi) \mapsto (\beta, X, z, \theta, Y)(A, \Phi),$$

defined for (A, Φ) close to $(\bar{A}, \bar{\Phi})$, such that

$$(\beta, X, z, \theta, Y)(\bar{A}, \bar{\Phi}) = (\alpha(\bar{X}), \bar{X}, \bar{z}, \bar{\theta}, \bar{Y}),$$

$(\beta, X, z)(A, \Phi)$ is the unique critical point near $(\alpha(\bar{X}), \bar{X}, \bar{z})$ of the perturbed problem $SAM(A, \Phi)$, and its corresponding Lagrange multiplier is $(\theta, Y)(A, \Phi)$. Furthermore, by smoothness, the second-order conditions will also hold at this point.

We now claim the point $X(A, \Phi)$ is our required nondegenerate critical point. Certainly all maps Φ close to $\bar{\Phi}$ are injective. Furthermore, at any matrix X close to $\bar{X}$ in the active manifold $\mathcal{M}$, it is easy to check Φ is transversal to $\mathcal{M}$ at X: in particular, this holds at $X(A, \Phi)$.

Note that $X(A, \Phi)$ is a strong minimizer of $\alpha|_{\mathcal{M}\cap(A+R(\Phi))}$, by the second-order conditions. In particular, it lies in the active manifold $\mathcal{M}$ and is close to $\bar{X}$, so it has the same active Jordan structure as $\bar{X}$, and hence has nonderogatory spectral abscissa. It remains to show $0 \in \operatorname{ri} \Phi^*\partial\alpha(X(A, \Phi))$.

Now, by analogy with (3.7), stationarity of the Lagrangian implies

$$\Phi^*(\hat{P}^*\Lambda(\hat{\theta})\hat{P}^{-*}) = 0, \qquad \text{where} \quad \hat{P} = P(X(A, \Phi)) \quad \text{and} \quad \hat{\theta} = \theta(A, \Phi):$$

we simply reduce the matrix $X(A, \Phi)$ to Arnold form (2.2) and apply Lemma 2.12 at $X(A, \Phi)$ in place of $\bar{X}$. By Theorem 2.15 (spectral abscissa subgradients) applied at $X(A, \Phi)$ we just need to show

$$\theta(A, \Phi) \in \operatorname{ri} \Theta_1^{\geq},$$

for (A, Φ) close to $(\bar{A}, \bar{\Phi})$. But this follows since

$$\theta(A, \Phi)) \in \Theta_1 = \operatorname{aff} \Theta_1^{\geq},$$

and as (A, Φ) approaches $(\bar{A}, \bar{\Phi})$ we know

$$\theta(A, \Phi) \to \bar{\theta} \in \mathrm{ri}\, \Theta_1^{\geq},$$

using our assumption that $0 \in \mathrm{ri}\, \Phi^* \partial\alpha(\bar{X})$ and Theorem 3.8 (active-critical points). This completes the proof. □

Our argument has interesting computational consequences. Under the assumptions of Theorem 4.1, the critical point of the perturbed problem $\inf_{A+R(\Phi)} \alpha$ is the unique matrix X close to $\bar{X}$ satisfying the system of equations

$$\begin{aligned} X &= A + \Phi(z), \\ PXP^{-1} &= \mathrm{Diag}(B, 0, 0, \ldots, 0) + \sum_{j=1}^{p} (J_{j1} + \lambda_j J_{j0}), \\ \mathrm{Re}\,\lambda_j &= \beta \qquad (j = 1, 2, \ldots, p), \\ \sum_{j=1}^{p} m_j \theta_{j0} &= 1, \\ \Phi^*(Y) &= 0, \\ Y &= P^* \Lambda(\theta) P^{-*}, \end{aligned}$$

for some z, P, B, λ, β, θ, and Y close to $\bar{z}$, $\bar{P}$, $\bar{B}$, $\bar{\lambda}$, $\alpha(\bar{X})$, $\bar{\theta}$, and $\bar{Y}$, respectively: the solution must be unique, with the exception of the variables P and B, by Theorem 3.8 (active critical points).

Theorem 4.1 (persistence of Jordan form) shows that the active Jordan form of a nondegenerate critical point X for a spectral abscissa minimization problem is *robust* under small perturbations to the problem. If we know this active Jordan form, and we can find an approximation to X, then we can apply an iterative solution technique to the system of equations in the preceding paragraph to refine our estimate of X.

The definition of a nondegenerate critical point is, in part, motivated by the second-order sufficient conditions for a local minimizer in nonlinear programming: we require the objective function to have positive-definite Hessian on a certain active manifold, and certain Lagrange multipliers have their correct signs. This suggests the following question:

Question 4.2. Must nondegenerate critical points of spectral abscissa minimization problems be local minimizers?

5. Examples

Our main result, Theorem 4.1 (persistence of active Jordan form), assumes the critical matrix $\bar{X}$ satisfies the four conditions of Definition 1.3 (nondegenerate critical points). Simple examples show three of these conditions are all, in general, necessary for the result to hold.

Example 5.1 (Necessity of Transversality). Consider the problem

$$\inf\left\{\alpha\begin{bmatrix} z & 1 \\ 0 & -z \end{bmatrix} : z \in \mathbf{R}\right\},$$

which clearly has global minimizer $\bar{z} = 0$, since $\bar{f} = |\cdot|$. The map $\bar{\Phi}\colon \mathbf{R} \to \mathbf{M^2}$, defined by $\bar{\Phi}(z) = \left[\begin{smallmatrix} z & 0 \\ 0 & -z \end{smallmatrix}\right]$, is *not* transversal to the active manifold $\mathcal{M}$ at the nonderogatory matrix $\bar{X} = J_2 = \bar{A}$, since by Corollary 2.14 (normal space) we know

$$N_{\mathcal{M}}(\bar{X}) = \begin{bmatrix} 0 & 0 \\ \mathbf{C} & 0 \end{bmatrix} \tag{5.2}$$

so, for example,

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \in N(\bar{\Phi}^*) \cap N_{\mathcal{M}}(\bar{X}).$$

Nonetheless, $\bar{X}$ is a strong minimizer of α on $\mathcal{M} \cap (\bar{A} + R(\bar{\Phi})) = \{\bar{X}\}$. Furthermore, we have

$$\bar{\Phi}^*\partial\alpha(\bar{X}) = \left\{\operatorname{Re}\operatorname{tr}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 \\ \theta & \frac{1}{2} \end{bmatrix} : \operatorname{Re}\theta \geq 0\right\} = \{0\},$$

so $0 \in \operatorname{ri} \bar{\Phi}^*\partial\alpha(\bar{X})$. (In this case, the chain rule $\partial\bar{f}(\bar{z}) = \bar{\Phi}^*\partial\alpha(\bar{X})$ fails due to the lack of transversality, so it is worth noting that we also have $0 \in \operatorname{ri}\partial\bar{f}(\bar{z})$.) Despite this, perturbing the problem may change the Jordan form of the corresponding critical point: for example, for all real $\varepsilon > 0$ the function

$$\alpha\begin{bmatrix} z & 1 \\ \varepsilon & -z \end{bmatrix} = \sqrt{\varepsilon + z^2}$$

has unique critical point $z = 0$, and the corresponding matrix $\left[\begin{smallmatrix} 0 & 1 \\ \varepsilon & 0 \end{smallmatrix}\right]$ has distinct eigenvalues.

Example 5.3 (Necessity of Strong Minimality). Consider the problem

$$\inf\{\alpha(1 + z \cdot 0) : z \in \mathbf{R}\}.$$

The identically zero map $\bar{\Phi}\colon \mathbf{R} \to \mathbf{C}$ is transversal to the active manifold $\mathcal{M}$ (which is just $\mathbf{C}$, locally) at the nonderogatory one-by-one matrix $\bar{X} = 1 = \bar{A}$, which corresponds to the critical point $\bar{z} = 0$. Furthermore, $0 \in \operatorname{ri}\bar{\Phi}^*\partial\alpha(\bar{X}) = \{0\}$. However, nearby problems like $\inf\alpha(1 + z \cdot \varepsilon)$ (for real nonzero ε) have no critical points. In this case it is the strong minimizer condition that fails.

Example 5.4 (Necessity of Strict First-Order Condition). Consider the problem

$$\inf\left\{\alpha\begin{bmatrix}0 & 1\\ z & 0\end{bmatrix}:\ z\in\mathbf{C}\right\}.$$

The map $\bar{\Phi}\colon \mathbf{R}\to\mathbf{M}^2$, defined by $\bar{\Phi}(z)=\begin{bmatrix}0 & 0\\ z & 0\end{bmatrix}$, is transversal to the active manifold at the nonderogatory matrix $\bar{X}=J_2=\bar{A}$ corresponding to $\bar{z}=0$, using (5.2). We have

$$\bar{\Phi}^*\partial\alpha(\bar{X})=\left\{\operatorname{Re}\operatorname{tr}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{2} & 0\\ \theta & \frac{1}{2}\end{bmatrix}:\ \operatorname{Re}\theta\geq 0\right\}=\mathbf{R}_+,$$

so $\bar{X}$ is a critical point, and is a strong minimizer of α on $\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))=\{\bar{X}\}$. However, for real $\varepsilon>0$, perturbing $\bar{\Phi}$ to $\Phi(z)=\begin{bmatrix}\varepsilon z & 0\\ z & \varepsilon z\end{bmatrix}$ results in a unique point in $\mathcal{M}\cap(\bar{A}+R(\Phi))$, namely $\bar{X}$, and now

$$\bar{\Phi}^*\partial\alpha(\bar{X})=\left\{\operatorname{Re}\operatorname{tr}\begin{bmatrix}\varepsilon & 1\\ 0 & \varepsilon\end{bmatrix}\begin{bmatrix}\frac{1}{2} & 0\\ \theta & \frac{1}{2}\end{bmatrix}:\ \operatorname{Re}\theta\geq 0\right\}=[\varepsilon,+\infty),$$

so $\bar{X}$ is no longer a critical point. In this case it is the failure of the strict first-order condition $0\in\operatorname{ri}\bar{\Phi}^*\partial\alpha(\bar{X})$ which causes our result to break down.

None of the examples above address the case where the critical matrix $\bar{X}$ has derogatory spectral abscissa. The lack of obvious examples suggests the following problem:

Question 5.5 (Derogatory Minimizers). Find a matrix X with derogatory spectral abscissa such that X is a strict local minimizer of the spectral abscissa on the affine set $A+R(\Phi)$ and the linear map Φ is transversal to the active manifold at X.

Consider, for example, the case of the n-by-n zero matrix $X=0$, with $n>1$. Clearly the active manifold $\mathcal{M}$ is the set of all small complex multiples of the identity matrix. Hence if the map Φ is transversal to $\mathcal{M}$ at 0, then its range has real codimension at most 2. Thus $R(\Phi)$ certainly contains a subspace of the form $S=\{Z\colon \operatorname{tr}KZ=0\}$ for some nonzero matrix K. If K is a multiple of the identity, then for all real ε the diagonal matrix $\operatorname{Diag}(\varepsilon i,-\varepsilon i,0,0,\ldots,0)$ lies in S and has spectral abscissa 0, so 0 is not a strict local minimizer. On the other hand, if K is not a multiple of the identity, then an elementary argument shows S contains a matrix Z with $\alpha(Z)<0$, so by considering εZ for real $\varepsilon>0$ we see again that 0 is not a local minimizer. In summary, the zero matrix can never solve Question 5.5.

We end with a simple example illustrating the main result.

Example 5.6. Consider the spectral abscissa minimization problem

$$\inf\left\{\alpha\begin{bmatrix}-1 & 0 & z_1\\ 0 & z_2 & 1\\ z_1^* & -z_2 & 0\end{bmatrix} :\ z\in\mathbf{C}^2\right\},$$

and we consider the feasible solution

$$\bar{X}=\begin{bmatrix}-1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}=\bar{A},$$

corresponding to $\bar{z}=0$. Clearly $\bar{X}$ has nonderogatory spectral abscissa 0. In our notation we have $n=3$, $p=1$, $m_0=1$, $m_1=2$, $\bar{\lambda}=0$, $\bar{P}=I$, and $\bar{B}=-1$. The space $\mathbf{E}$ is $\mathbf{C}^2$ and the map $\bar{\Phi}\colon\ \mathbf{C}^2\to\mathbf{M}^3$ is defined by

$$\Phi(z)=\begin{bmatrix}0 & 0 & z_1\\ 0 & z_2 & 0\\ z_1^* & -z_2 & 0\end{bmatrix}.$$

It is clearly injective, with adjoint $\bar{\Phi}^*\colon\ \mathbf{M}^3\to\mathbf{C}^2$ defined by

$$\bar{\Phi}^*(Y)=(y_{13}+y_{31}^*,\,y_{22}-y_{32}).$$

The space Θ is $\mathbf{R}\times\mathbf{C}$ and we have

$$\Theta_0=\{(\theta_{10},\theta_{11})\colon\ \theta_{10}=0\},$$

and

$$\Theta_1^{\geq}=\{(\theta_{10},\theta_{11})\colon\ \theta_{10}=\tfrac{1}{2},\ \operatorname{Re}\theta_{11}\geq 0\},$$

so

$$N_{\mathcal{M}}(\bar{X})=\left\{\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & \theta_{11} & 0\end{bmatrix} :\ \theta_{11}\in\mathbf{C}\right\}.$$

Thus $\bar{\Phi}$ is transversal to $\mathcal{M}$ at $\bar{X}$. Furthermore,

$$\partial\alpha(\bar{X})=\left\{\begin{bmatrix}0 & 0 & 0\\ 0 & \frac{1}{2} & 0\\ 0 & \theta_{11} & \frac{1}{2}\end{bmatrix} :\ \operatorname{Re}\theta_{11}\geq 0\right\},$$

so the Lagrange multipliers are given by

$$(\bar{\theta}_{10},\bar{\theta}_{11})=(\tfrac{1}{2},\tfrac{1}{2})\in\operatorname{ri}\Theta_1^{\geq}\qquad\text{and}\qquad\bar{Y}=\begin{bmatrix}0 & 0 & 0\\ 0 & \frac{1}{2} & 0\\ 0 & \frac{1}{2} & \frac{1}{2}\end{bmatrix}.$$

Hence, by Theorem 3.8 (active critical points), $0 \in \text{ri}\,\bar{\Phi}^*\partial\alpha(\bar{X})$. (Alternatively, direct calculation shows $\bar{\Phi}^*\partial\alpha(\bar{X}) = \{(0, \frac{1}{2} - \theta_{11})\colon \text{Re}\,\theta_{11} \geq 0\}$.)

It remains to check that $\bar{X}$ is a strong minimizer of $\alpha|_{\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))}$. Checking this directly by Theorem 3.8 involves computing the second derivatives of the functions in the Arnold form: we avoid this by a direct calculation.

By transversality, the tangent space to the feasible region is

$$T_{\mathcal{M}\cap(\bar{A}+R(\bar{\Phi}))}(\bar{X}) = T_{\mathcal{M}}(\bar{X}) \cap R(\bar{\Phi}) = \left\{ \begin{bmatrix} 0 & 0 & z_1 \\ 0 & 0 & 0 \\ z_1^* & 0 & 0 \end{bmatrix} : z_1 \in \mathbf{C} \right\}.$$

Hence if the matrix X is close to $\bar{X}$ in $\mathcal{M} \cap (\bar{A} + R(\bar{\Phi}))$, then

$$X = \begin{bmatrix} -1 & 0 & z_1 \\ 0 & z_2 & 1 \\ z_1^* & -z_2 & 0 \end{bmatrix} \qquad \text{with} \quad z_2 = o(z_1).$$

This matrix has characteristic polynomial

$$\lambda^3 + (1 - z_2)\lambda^2 - |z_1|^2\lambda + z_2(1 + |z_1|^2).$$

The derivative of this polynomial,

$$3\lambda^2 + 2(1 - z_2)\lambda - |z_1|^2,$$

has roots

$$\tfrac{1}{3}((z_2 - 1) \pm ((1 - z_2)^2 + 3|z_1|^2)^{1/2}).$$

The root with largest real part is

$$\frac{1 - z_2}{3}\left(\left(1 + \frac{3|z_1|^2}{(1 - z_2)^2}\right)^{1/2} - 1\right) \sim \frac{|z_1|^2}{2}$$

for small z. By Gauss's theorem, the derivative has roots in the convex hull of the roots of the characteristic polynomial, so

$$\alpha(X) \geq \frac{|z_1|^2}{3} \sim \frac{\|z\|^2}{3}$$

for small z. Thus $\bar{X}$ is a strong minimizer on the tangent space to the feasible region, as required.

In conclusion, $\bar{X}$ is a nondegenerate critical point for the original spectral abscissa minimization problem. Hence, by Theorem 4.1 (persistence of active Jordan form), for any small perturbation to the problem there is a nondegenerate critical point close to $\bar{X}$ with the same active Jordan structure.

Acknowledgments

The research of J. V. Burke was supported by NSF; the research of A. S. Lewis was supported by NSERC; and the research of M. L. Overton was supported by NSF, DOE.

References

[1] V. I. Arnold, On matrices depending on parameters, *Russian Math. Surveys*, **26** (1971), 29–43.

[2] V. Blondel, M. Gevers, and A. Lindquist, Survey on the state of systems and control, *European J. Control*, **1** (1995), 5–23.

[3] V. Blondel and J. N. Tsitsiklis, NP-hardness of some linear control design problems, *SIAM J. Control Optim.*, **35** (1997), 2118–2127.

[4] J. V. Burke, A. S. Lewis, and M. L. Overton, Optimizing matrix stability, *Proc. Amer. Math. Soc.*, (2000), to appear.

[5] J. V. Burke and M. L. Overton, Variational analysis of non-Lipschitz spectral functions, *Math. Programming*, (2001), to appear.

[6] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, 1990.

[7] L. El Ghaoui and S.-I. Niculescu (eds.), *Advances in Linear Matrix Inequality Methods in Control*, SIAM, Philadelphia, 2000.

[8] A. A. Mailybaev, Transformation of families of matrices to normal form and its application to stability theory, *SIAM J. Matrix Anal. Appl.*, **21** (1999), 396–417.

[9] B. T. Polyak and P. S. Shcherbakov, Numerical search of stable or unstable element in matrix or polynomial families: A unified approach to robustness analysis and stabilization, Technical report, Institute of Control Science, Moscow, 1999.

[10] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.