OPTIMIZING MATRIX STABILITY

J.V. Burke[*], **A.S. Lewis**[†], **and M.L. Overton**[‡]

September 28, 1999

Key words: eigenvalue optimization, spectral abscissa, nonsmooth analysis, Jordan form
AMS 1991 Subject Classification:
Primary: 15A42, 90C30
Secondary: 65F15, 49K30

Abstract

Given an affine subspace of square matrices, we consider the problem of minimizing the spectral abscissa (the largest real part of an eigenvalue). We give an example whose optimal solution has Jordan form consisting of a single Jordan block, and we show, using non-lipschitz variational analysis, that this behaviour persists under arbitrary small perturbations to the example. Thus although matrices with nontrivial Jordan structure are rare in the space of all matrices, they appear naturally in spectral abscissa minimization.

1 Introduction

The *spectral abscissa* of a square matrix is the largest of the real parts of its eigenvalues. Our interest here is in minimizing the spectral abscissa of a parametrized matrix.

Its intrinsic mathematical interest aside, this problem is fundamental in control theory, where the spectral abscissa of a matrix A determines the stability of the dynamical system $u' = Au$. In practice, considerations

[*]Department of Mathematics, University of Washington, Seattle, WA 98195, U.S.A. Email: `burke@math.washington.edu`. Research supported by NSF.

[†]Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. Email: `aslewis@math.uwaterloo.ca`. Research supported by NSERC.

[‡]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, U.S.A. Email: `overton@cs.nyu.edu`. Research supported by NSF, DOE.

of transient behaviour and the effects of forcing terms or nonlinearity may complicate stability questions ([4] gives some interesting infinite-dimensional illustrations, for example). Nonetheless, understanding the asymptotic behaviour of the pure, homogeneous linear model is crucial.

Consider for instance the damped linear oscillator equation

$$w'' + \mu w' + w = 0,$$

where μ is a real parameter. By writing $u = (w\ w')^T$ we obtain the system

$$u' = \begin{pmatrix} 0 & 1 \\ -1 & -\mu \end{pmatrix} u.$$

The spectral abscissa of the defining matrix is minimized when $\mu = 2$, and the optimal matrix has Jordan form

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

The appearance of a nontrivial Jordan block is at first sight a little surprising, since such matrices all lie in a manifold of smaller dimension than that of the underlying space. Simple computational experiments suggest this is not an isolated phenomenon but happens quite commonly: minimizing the spectral abscissa over an affine manifold of matrices seems to encourage the appearance of matrices with nontrivial (and hence nongeneric) Jordan structure. Our aim in this paper is to make this observation more precise.

Just as the stability of continuous-time dynamical systems is closely related to the spectral abscissa, the discrete-time case corresponds to the *spectral radius* of a square matrix (the largest of the moduli of its eigenvalues). Our discussion of the spectral abscissa here has a complete analogue for the spectral radius, which we do not pursue.

We begin by presenting a particular example of a spectral abscissa minimization problem whose solution has a large Jordan block. Specifically, we present an $(n-1)$-dimensional affine subspace of $n \times n$ matrices on which the spectral abscissa is uniquely minimized at a matrix consisting of a single Jordan block: our proof is direct and classical.

On a general affine subspace, verifying that a given matrix minimizes the spectral abscissa is challenging: the relevant tools, combining matrix and nonsmooth analysis, have begun to appear only very recently. Two of the authors [2] have studied the spectral abscissa's subdifferential (in the sense of [3]). Using this work we verify that the optimal solution in our particular example is a 'sharp' local minimizer: that is, the function value grows at least linearly near the minimum.

We next consider the effect of perturbing our example. Using work of Arnold [1] on canonical forms for parametrized matrices, we show that any affine subspace sufficiently close to the original subspace contains a unique matrix close to the original solution whose Jordan form consists of a single block: furthermore, variational calculus shows this matrix is also a sharp local minimizer of the spectral abscissa over the subspace. In summary, the large Jordan block in the solution to our example *persists* under arbitrary small perturbations.

Our example shows that, although matrices with large Jordan blocks are rare in the space of all matrices, they appear not as singularities in spectral abscissa minimization but instead rather naturally. Given the practical importance of these problems, this emphasizes the importance of subdifferential analysis for matrices with nontrivial Jordan structure. Our approach is also a striking illustration of the power of modern variational analysis, as expounded in [3] for example, to attack concrete nonconvex nonlipschitz problems of great significance.

2 The basic example

The example we describe in this section is central to the paper. We denote the set of $n \times n$ complex matrices by $\mathbf{M}^n$. The *spectral abscissa* $\alpha : \mathbf{M}^n \to \mathbf{R}$ maps matrices to the largest real part of their eigenvalues.

Theorem 2.1 *All nonzero vectors x in $\mathbf{C}^{n-1}$ satisfy the inequality*

$$\alpha \begin{pmatrix} -x_1 & 1 & 0 & \dots & 0 \\ x_1 & 0 & 1 & \dots & 0 \\ x_2 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ x_{n-1} & 0 & 0 & \dots & 0 \end{pmatrix} > 0.$$

Thus $x = 0$ is a strict global minimizer of the function on the left-hand-side.

Proof A standard induction argument shows the characteristic polynomial of the matrix above is given by

$$p_x(\lambda) = \lambda^n + x_1 \lambda^{n-1} - \sum_{j=2}^{n} x_{j-1} \lambda^{n-j}.$$

Suppose all the roots of p_x have nonpositive real parts. By a classical result of Gauss, the derivatives $p_x^{(1)}, p_x^{(2)}, p_x^{(3)}, \ldots, p_x^{(n-2)}$ have all their roots in the convex hull of the roots of p_x, and hence all these roots have nonpositive real parts.

Denote the roots of the quadratic

$$p_x^{(n-2)}(\lambda) = \frac{n!}{2}\lambda^2 + (n-1)!x_1\lambda - (n-2)!x_1$$

by λ_1 and λ_2. If x_1 is nonzero then we know

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_2} = \frac{\lambda_1 + \lambda_2}{\lambda_1\lambda_2} = n - 1,$$

which implies at least one of λ_1^{-1} and λ_2^{-1} has strictly positive real part, whence so does λ_1 or λ_2. Thus $x_1 = 0$.

Now we proceed by induction. For $j = 2, 3, \ldots, n$, assuming $x_1 = x_2 = x_3 = \ldots = x_{j-1} = 0$, we deduce $x_j = 0$ since the roots of the polynomial

$$p_x^{(n-j-1)}(\lambda) = \frac{n!}{(j+1)!}\lambda^{j+1} - (n-j-1)!x_j$$

all have nonpositive real part. Hence we see $x = 0$. ♣

3 The Jordan manifold

Our argument depends on the analytic structure of the set $\mathcal{M}$ of matrices in $\mathbf{M}^n$ whose Jordan form consists of a single block. If we define the *Jordan matrix* J in $\mathbf{M}^n$ by

$$J = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix},$$

then

$$\mathcal{M} = \{S^{-1}JS + \delta I : \delta \in \mathbf{C},\ S \in GL_n\},$$

where GL_n denotes the set of invertible matrices in $\mathbf{M}^n$. We call this set $\mathcal{M}$ the *Jordan manifold*. We summarize the properties we need in the next result. We make $\mathbf{M}^n$ into a Euclidean space (that is, a real inner product space) by endowing it with the obvious vector space operations and the inner product $\langle X, Y\rangle = \operatorname{Re}\operatorname{tr} X^*Y$.

Theorem 3.1 (Jordan manifold) *The Jordan manifold $\mathcal{M}$ is a submanifold of $\mathbf{M}^n$ of complex codimension $n-1$. The tangent space to $\mathcal{M}$ at the matrix J is given by*

$$T_J(\mathcal{M}) = \{Z \in \mathbf{M}^n : \langle (J^j)^*, Z\rangle = 0 \quad (j = 1, 2, 3, \ldots, n-1)\}.$$

If the matrix X is close to J in $\mathcal{M}$ then

$$X = S^{-1}JS + \delta I \tag{3.2}$$

for some small complex δ and some matrix S close to I.

Proof The manifold and tangent space structure may be found in [1]. If X is close to J in $\mathcal{M}$ then X is given by equation (3.2) and its unique eigenvalue δ must be close to the unique eigenvalue of J, namely 0.

By [1, §3.4] there is a neighbourhood Ω of J in $\mathbf{M}^n$ and holomorphic functions $f : \Omega \to \mathbf{C}^n$ and $T : \Omega \to GL_n$ satisfying the equations $f(J) = 0$, $T(J) = I$, and

$$Y = T(Y)^{-1} \begin{pmatrix} f_1(Y) & 1 & 0 & \ldots & 0 \\ f_2(Y) & 0 & 1 & \ldots & 0 \\ f_3(Y) & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ f_n(Y) & 0 & 0 & \ldots & 0 \end{pmatrix} T(Y), \quad \text{for all } Y \in \Omega.$$

The characteristic polynomial of Y is $\lambda^n - \sum_{j=1}^n f_j(Y)\lambda^{n-j}$. Since $X - \delta I$ and J are similar, their characteristic polynomials coincide, so $f(X - \delta I) = 0$, and X is close to J so $T(X - \delta I)$ is close to I, as required. $\diamondsuit$

4 Subdifferentials and sharp minima

Given a Euclidean space $\mathbf{E}$ and a function $f : \mathbf{E} \to [-\infty, \infty]$, the *regular subdifferential* of f at a point x in $\mathbf{E}$ at which f is finite is the set

$$\hat{\partial} f(x) = \{y \in \mathbf{E} : f(x+z) - f(x) \geq \langle y, z\rangle + o(z) \ \text{ for small } z \in \mathbf{E}\}.$$

If f is infinite at x, define $\hat{\partial} f(x) = \emptyset$. The *subdifferential* $\partial f(x)$ consists of those elements y of $\mathbf{E}$ for which there exists a sequence of points x_r approaching x in $\mathbf{E}$ with $f(x_r)$ approaching $f(x)$ and a sequence of elements $y_r \in \hat{\partial} f(x_r)$ approaching y. If instead we have $t_r y_r$ approaching y

for some sequence of reals t_r decreasing to 0 we obtain the *horizon subdifferential* $\partial^\infty f(x)$: by definition $0 \in \partial^\infty f(x)$, and if f is infinite at x, define $\partial^\infty f(x) = \{0\}$. We say f is *subdifferentially regular* at x if it is locally lower semicontinuous, the subdifferential is nonempty and coincides with the regular subdifferential, and the horizon subdifferential coincides with the recession cone of the regular subdifferential (where the *recession cone* of a nonempty closed convex set C is the set of vectors d satisfying $c + td \in C$ for all vectors c in C and real $t \geq 0$). For more details, see [3].

For example, consider the spectral abscissa: the next result is due to [2].

Theorem 4.1 (Spectral abscissa subdifferential) *The spectral abscissa α is subdifferentially regular at the Jordan matrix J, with subdifferential given by*

$$\partial\alpha(J) = \left\{ \frac{1}{n} I + \sum_{j=1}^{n-1} \theta_j (J^j)^* : \theta \in \mathbf{C}^{n-1},\ \mathrm{Re}\,\theta_1 \geq 0 \right\}.$$

Our interest in subdifferentials stems from their use in optimality conditions. It is elementary to check, for example, that if a point $\bar{x}$ in $\mathbf{E}$ is a local minimizer of a function f then it must satisfy the 'necessary' condition $0 \in \hat{\partial} f(\bar{x})$. To obtain a 'sufficient' condition we need a stronger assumption. We call a local minimizer $\bar{x}$ of f *sharp* if $f(\bar{x})$ is finite and there exists a real $\delta > 0$ such that

$$(4.2) \qquad f(\bar{x} + z) - f(\bar{x}) \geq \delta \|z\| \quad \text{for all small } z \in \mathbf{E}.$$

Clearly this inequality implies $\delta \mathbf{B} \subset \hat{\partial} f(\bar{x})$, where $\mathbf{B}$ denotes the closed unit ball in $\mathbf{E}$. Conversely, for any real $\delta' > 0$, the inclusion $\delta' \mathbf{B} \subset \hat{\partial} f(\bar{x})$ implies inequality (4.2) holds for all real $\delta < \delta'$. Hence we have the following convenient (if strong) sufficient condition.

Proposition 4.3 (Sharp minima) *A point $\bar{x}$ is a sharp local minimizer of a function f if and only if $0 \in \mathrm{int}\, \hat{\partial} f(\bar{x})$.*

To analyze linearly parametrized examples like that of Theorem 2.1, we need a chain rule. The following suffices for us. It concerns another Euclidean space $\mathbf{Y}$, a linear map $A : \mathbf{E} \to \mathbf{Y}$, and its adjoint $A^* : \mathbf{Y} \to \mathbf{E}$.

Lemma 4.4 (Chain rule) *If the function f is subdifferentially regular at 0 and the linear map A satisfies the condition*

$$A^* y = 0 \text{ and } y \in \partial^\infty f(0) \Rightarrow y = 0,$$

then the composite function $f \circ A$ is subdifferentially regular at 0 *with subdifferential*

$$\partial(f \circ A)(0) = A^* \partial f(0).$$

Proof This follows from [3, Thm 10.6]. ♡

We can now see that the global minimizer we described in Theorem 2.1 is also a sharp local minimizer.

Example 4.5 Define a linear map $A_0 : \mathbf{C}^{n-1} \to \mathbf{M}^n$ by

$$x \in \mathbf{C}^{n-1} \mapsto \begin{pmatrix} -x_1 & 0 & 0 & \dots & 0 \\ x_1 & 0 & 0 & \dots & 0 \\ x_2 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ x_{n-1} & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{4.6}$$

The adjoint $A_0^* : \mathbf{M}^n \to \mathbf{C}^{n-1}$ is given by

$$(y_{rs}) \in \mathbf{M}^n \mapsto (y_{21} - y_{11}, y_{31}, y_{41}, \dots, y_{n1})^T.$$

Define the function $f : \mathbf{M}^n \to \mathbf{R}$ by

$$X \in \mathbf{M}^n \mapsto \alpha(J + X).$$

By Theorem 4.1 (Spectral abscissa subdifferential) we know f is subdifferentially regular at 0 with subdifferential

$$\partial f(0) = \left\{ \frac{1}{n} I + \sum_{j=1}^{n-1} \theta_j (J^j)^* : \theta \in \mathbf{C}^{n-1},\ \mathrm{Re}\, \theta_1 \geq 0 \right\},$$

and horizon subdifferential

$$\partial^\infty f(0) = (\partial f(0))^\infty = \left\{ \sum_{j=1}^{n-1} \theta_j (J^j)^* : \theta \in \mathbf{C}^{n-1},\ \mathrm{Re}\, \theta_1 \geq 0 \right\}.$$

For any vector θ in $\mathbf{C}^{n-1}$ we have

$$A_0^* \sum_{j=1}^{n-1} \theta_j (J^j)^* = \theta,$$

so the assumptions for the chain rule (Lemma 4.4) hold. We deduce

$$\begin{aligned} &\hat{\partial}(f \circ A_0)(0) = \partial(f \circ A_0)(0) = A_0^* \partial f(0) \\ &= \{(\theta_1 - 1/n, \theta_2, \theta_3, \ldots, \theta_{n-1})^T : \theta \in \mathbf{C}^{n-1},\ \mathrm{Re}\,\theta_1 \geq 0\}, \end{aligned}$$

and since the interior of this set contains 0, we see 0 is a sharp local minimizer for the example of Theorem 2.1.

5 Perturbation

We are interested in *linearly parametrized spectral abscissa minimization problems*

$$\inf_{x \in \mathbf{C}^k} \alpha(D + Ax),$$

for a given matrix D in $\mathbf{M}^n$ and a linear map $A : \mathbf{C}^k \to \mathbf{M}^n$. In the example in Theorem 2.1, D is the Jordan matrix J and A is the map A_0 defined by (4.6). In this case we showed $\bar{x} = 0$ is a sharp local minimizer: the corresponding matrix $D + A\bar{x}$ clearly has Jordan form consisting of a single Jordan block. Our aim in this final section is to show that this behaviour persists for arbitrary choices of D close to J and A close to A_0.

We start with a simple lemma.

Lemma 5.1 *Given a linear map $A : \mathbf{E} \to \mathbf{Y}$ and a convex set $C \subset \mathbf{E}$, if 0 lies in the interior of AC then it also lies in the interior of $\tilde{A}C$ for all linear maps $\tilde{A}$ close to A.*

Proof If the result fails then for some sequence of maps A_r approaching A we can separate 0 from $A_r C$. Hence there are unit vectors y_r in $\mathbf{Y}$ satisfying

$$\langle y_r, A_r x \rangle \geq 0 \ \text{ for all } x \in C,\ r = 1, 2, 3, \ldots.$$

We can assume y_r approaches a unit vector y in $\mathbf{Y}$, and then we have the contradiction

$$\langle y, Ax \rangle = \lim_r \langle y_r, A_r x \rangle \geq 0 \ \text{ for all } x \in C,$$

which completes the proof. ♠

Returning to the context of the chain rule (Lemma 4.4), we next show sharp local minima persist under arbitrary small perturbations of linearly parametrized problems.

Theorem 5.2 (Sharp minima persist) *Suppose the function f is subdifferentially regular at 0 and the linear map A satisfies the condition*

$$A^* y = 0 \text{ and } y \in \partial^\infty f(0) \;\Rightarrow\; y = 0.$$

If 0 is a sharp local minimizer of the composite function $f \circ A$ then it is also a sharp local minimizer of $f \circ \tilde{A}$ for all linear maps $\tilde{A}$ close to A.

Proof We first claim, for all $\tilde{A}$ close to A, the property

$$\tilde{A}^* y = 0 \text{ and } y \in \partial^\infty f(0) \;\Rightarrow\; y = 0.$$

If this fails, there is a sequence of linear maps A_r approaching A and a sequence of nonzero horizon subgradients y_r in $\partial^\infty f(0)$ satisfying $A_r^* y_r = 0$. Since $\partial^\infty f(0)$ is a closed cone, we can assume (y_r) is a sequence of unit vectors converging to another unit vector y in $\partial^\infty f(0)$, but then

$$A^* y = \lim_r A_r^* y_r = 0,$$

which is a contradiction.

By the chain rule and Proposition 4.5 (Sharp minima) we know

$$0 \in \operatorname{int} \partial(f \circ A)(0) = \operatorname{int} A^* \partial f(0).$$

The previous lemma now shows $0 \in \operatorname{int} \tilde{A}^* \partial f(0)$ for all $\tilde{A}$ close to A, so the chain rule applies again to show $0 \in \operatorname{int} \partial(f \circ \tilde{A})(0)$, which completes the proof. △

Theorem 5.3 (Perturbation) *If the matrix $D \in \mathbf{M}^n$ is close to the Jordan matrix J and the linear map $A : \mathbf{C}^{n-1} \to \mathbf{M}^n$ is close to the linear map A_0 defined by (4.6) then there is a unique point $\bar{x}$ close to 0 in $\mathbf{C}^{n-1}$ with $D + A\bar{x}$ in the Jordan manifold $\mathcal{M}$. Furthermore, $\bar{x}$ is a sharp local minimizer for the linearly-constrained spectral abscissa minimization problem*

$$\inf_{x \in \mathbf{C}^{n-1}} \alpha(D + Ax).$$

Proof The affine manifold $J + A_0 \mathbf{C}^{n-1}$ has dimension $n-1$, the Jordan manifold $\mathcal{M}$ has codimension $n-1$, and it is easy to check, by Theorem 3.1 (Jordan manifold), the equivalence

$$A_0 x \in T_J(\mathcal{M}) \quad \Leftrightarrow \quad x = 0.$$

Hence the map $x \mapsto J + A_0 x$ intersects $\mathcal{M}$ transversally at $x = 0$. The first claim now follows from the inverse function theorem.

According to Theorem 3.1, we know $D + A\bar{x} = S^{-1}JS + \delta I$ for some small complex δ and some matrix S close to I, so we can rewrite our problem

$$\inf_{x \in \mathbf{C}^{n-1}} \alpha(S^{-1}JS + \delta I + A(x - \bar{x})),$$

or equivalently (with the change of variables $z = x - \bar{x}$)

$$\inf_{z \in \mathbf{C}^{n-1}} \alpha(J + S(Az)S^{-1}). \tag{5.4}$$

As we observed in Example 4.5, the unperturbed problem

$$\inf_{z \in \mathbf{C}^{n-1}} \alpha(J + A_0 z)$$

has a sharp local minimum at 0. The linear map $z \in \mathbf{C}^{n-1} \mapsto S(Az)S^{-1}$ is close to A_0. This, together with our calculations in Example 4.5, shows Theorem 5.2 (Sharp minima persist) applies, so 0 is also a sharp local minimizer of problem (5.4), which proves $\bar{x}$ is a sharp local minimizer of the original perturbed problem, as required. $\bigtriangledown$

The unperturbed problem in the above result in fact has a strict *global* minimum at 0, by Theorem 2.1. However, simple examples show the sharp *local* minimizers for the perturbed problems may not be global minimizers.

References

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

[2] J.V. Burke and M.L. Overton. Variational analysis of non-Lipschitz spectral functions, September 1999. Submitted to *Mathematical Programming.*

[3] R.T. Rockafellar and R.J.-B. Wets. *Variational analysis.* Springer, Berlin, 1998.

[4] L.N. Trefethen. Pseudospectra of linear operators. *SIAM Review*, 39:383–406, 1997.