Computational Optimization and Applications 13, 13–23 (1999)

The Clarke and Michel-Penot Subdifferentials of the Eigenvalues of a Symmetric Matrix

J.-B. HIRIART-URRUTY jbhu@cict.fr
LAO, U.F.R. Mathématiques, Informatique, Gestion, Université Paul Sabatier, 118, route de Narbonne, F-31062 Toulouse Cedex, France

A.S. LEWIS* aslewis@orion.uwaterloo.ca
Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. http://orion.uwaterloo.ca/∼aslewis

Received October 1, 1997; Accepted February 9, 1998

Abstract. We calculate the Clarke and Michel-Penot subdifferentials of the function which maps a symmetric matrix to its mth largest eigenvalue. We show these two subdifferentials coincide, and are identical for all choices of index m corresponding to equal eigenvalues. Our approach is via the generalized directional derivatives of the eigenvalue function, thereby completing earlier studies on the classical directional derivative.

Keywords: eigenvalue optimization, nonsmooth analysis, Michel-Penot subdifferential, Clarke subdifferential, generalized derivative

Dedication: In honour of Professor Olvi Mangasarian

1. Introduction

Eigenvalue optimization is an important testing-ground for nonsmooth optimization theory. Optimization problems involving the eigenvalues of a real symmetric matrix arise in many applications, from engineering design to graph-partitioning: two extensive surveys are [15, 22]. More general nonsmooth optimization problems are also common in engineering (see for example [20]).

If we denote the mth largest eigenvalue (counted with multiplicity) of a symmetric matrix A by $\lambda_m(A)$, then the function λ_m may be nonsmooth. Generalized subdifferentials are therefore good tools for any variational study of eigenvalues. In this paper we calculate the subdifferentials of λ_m in the sense of both Clarke [3] and Michel-Penot [16]. We find they coincide, and are equal for all values of the index m corresponding to the same eigenvalue.

Our calculation of the Michel-Penot subdifferential is new. The result for the Clarke subdifferential has been proved previously, by various approaches. One approach uses the characterization of the Clarke subdifferential as the convex hull of cluster points of

*Research partially supported by the Natural Sciences and Engineering Research Council of Canada.

gradients of the function at a local set of points of full measure [6, 13]. Another approach first calculates the 'approximate' subdifferential (of Ioffe-Kruger-Mordukhovich) and then constructs its convex hull [14].

By contrast to these techniques, our approach is more elementary and direct, and emphasizes a pleasing parallel between the Michel-Penot and Clarke calculations. Our starting point is a known expression for the usual directional derivative $\lambda_m'(A; G)$. We then simply regularize this derivative in two different ways: allowing perturbations of the base matrix A leads to the Clarke result, whereas allowing (suitable) perturbations of the direction matrix G gives the Michel-Penot version.

2. Subdifferentials

Consider an open subset Ω of $\mathbf{R}^n$ and a function $f : \Omega \to \mathbf{R}$. Given a point x in Ω, the (*usual* or *radial*) *directional derivative* of f at x is the positively homogeneous function

$$d \in \mathbf{R}^n \mapsto f'(x; d) = \lim_{t \to 0+} \frac{f(x + td) - f(x)}{t},$$

when this function exists—in this case we say f is *directionally differentiable* at x. Assuming, henceforth, f is Lipschitz near x, we can write the above function in the form

$$\lim_{t \to 0+, v \to d} \frac{f(x + tv) - f(x)}{t},$$

sometimes called the 'tangential directional derivative' of f at x (in the direction d).

The *Clarke directional derivative* is the function

$$d \in \mathbf{R}^n \mapsto f^\circ(x; d) = \limsup_{t \to 0+, y \to x} \frac{f(y + td) - f(y)}{t}.$$

This function is *sublinear*, by which we mean convex and positively homogeneous. When it coincides with the usual directional derivative we say the function f is *Clarke regular* (or 'strictly tangentially convex') at the point x. The *Clarke subdifferential* of f at x is the nonempty compact convex set

$$\partial^{\mathrm{cl}} f(x) = \{s \in \mathbf{R}^n : \langle s, d\rangle \le f^\circ(x; d) \text{ for all } d \text{ in } \mathbf{R}^n\}$$

(see [3]).

When the function f is directionally differentiable throughout the set Ω, we know, for all points x in Ω and directions d in $\mathbf{R}^n$,

$$f^\circ(x; d) = \limsup_{y \to x} f'(y; d) \tag{2.1}$$

(by [10, Section 2.2]). Thus the Clarke directional derivative is a 'regularized' version of the usual directional derivative, which explains the good topological behaviour of the map $x \mapsto f^\circ(x; d)$.

In general, formula (2.1) may have little practical use, since even when we know the usual directional derivative explicitly, taking the lim sup may be intractable. Nevertheless, this formula will be useful for eigenvalues.

There is an alternative to the Clarke directional derivative and subdifferential, due to Michel-Penot [16, 17]. The *Michel-Penot directional derivative* of the function f at the point x is the sublinear function

$$d \in \mathbf{R}^n \mapsto f^{\Diamond}(x; d) = \sup_{y \in \mathbf{R}^n} \limsup_{t \to 0+} \frac{f(x + ty + td) - f(x + td)}{t},$$

and the *Michel-Penot subdifferential* of f at x is the set

$$\partial^{\Diamond} f(x) = \{s \in \mathbf{R}^n : \langle s, d \rangle \leq f^{\Diamond}(x; d) \text{ for all } d \text{ in } \mathbf{R}^n\}.$$

If f is differentiable at x then $\partial^{\Diamond} f(x) = \{\nabla f(x)\}$ (by contrast with the Clarke version), but on the other hand the multifunction $\partial^{\Diamond} f(\cdot)$ does not enjoy the various continuity properties of $\partial^{\mathrm{cl}} f(\cdot)$.

Clearly in general we have

$$f'(x; \cdot) \leq f^{\Diamond}(x; \cdot) \leq f^{\circ}(x; \cdot),$$

so $\partial^{\Diamond} f(x) \subset \partial^{\mathrm{cl}} f(x)$. If the function f is positively homogeneous then it is easy to check

$$f^{\circ}(0; d) = f^{\Diamond}(0; d) = \sup_{y \in \mathbf{R}^n} \{f(y + d) - f(y)\} \tag{2.2}$$

for any direction d in $\mathbf{R}^n$ (cf. [10]), and hence $\partial^{\Diamond} f(0) = \partial^{\mathrm{cl}} f(0)$. Finally, note $f'(x; \cdot) = f^{\Diamond}(x; \cdot)$ if and only if the usual directional derivative exists and is convex (cf. [16]).

Sublinear majorants

We have seen that the Michel-Penot and Clarke directional derivatives are sublinear majorants of the ordinary directional derivative. In general, any positively homogeneous Lipschitz function $p : \mathbf{R}^n \to \mathbf{R}$ has a 'best' sublinear majorant ${}^*p : \mathbf{R}^n \to \mathbf{R}$, defined, for any direction d in $\mathbf{R}^n$, by

$${}^*p(d) = \sup_{y \in \mathbf{R}^n} \{p(y + d) - p(y)\}.$$

Clearly ${}^*p$ is a sublinear majorant of p, and is the smallest such function h satisfying the condition

$$p(y + d) \leq p(y) + h(d) \quad \text{for all } y \text{ and } d \text{ in } \mathbf{R}^n.$$

Equation (2.2) shows that ${}^*p$ is just the Clarke (or Michel-Penot) directional derivative of p at the origin. See [10] and the references therein for more on this operation, as well as [7, 8, 19]. If the function f is directionally differentiable at the point x then

$$f^{\diamond}(x; d) = \sup_{y \in \mathbf{R}^n} \{ f'(x; d + y) - f'(x; y) \}, \tag{2.3}$$

for all directions d in $\mathbf{R}^n$ ([17, Lemma 1.3]). Thus the Michel-Penot directional derivative is the 'best' sublinear majorant of the usual directional derivative:

$$f^{\diamond}(x; \cdot) = {}^*(f'(x; \cdot)).$$

If, on the other hand, the usual directional derivative fails to exist, one has to rely directly on the definition. As we shall see, this is not the case for the eigenvalue functions we consider.

3. Eigenvalues

We denote the inner product space of $n \times n$ real symmetric matrices by $\mathbf{S}^n$, where the inner product of two matrices A and B in $\mathbf{S}^n$ is $\langle A, B \rangle = \operatorname{tr} AB$. We write the mth largest eigenvalue of A (counted with multiplicity) as $\lambda_m(A)$ (for $m = 1, 2, \ldots, n$). Thus the function λ_m is positively homogeneous.

It is well-known that the function $\sigma_m = \sum_1^m \lambda_i$ is sublinear: indeed, it is the support function of the set

$$\{C \in \mathbf{S}^n : C \text{ and } I - C \text{ are positive semidefinite, } \operatorname{tr} C = m\} \tag{3.1}$$

(see, for example, [11] or [18]). Thus λ_1 is convex, λ_n is concave, and λ_m is the difference of two finite convex functions, σ_m and σ_{m-1}, for $1 < m < n$. In particular, λ_m is Lipschitz. (In fact the Lipschitz constant is 1.)

Associated with a fixed matrix A in $\mathbf{S}^n$ and a fixed index m are two other indices, $\hat{m}$ and $\bar{m}$, defined by

$$\hat{m} = \min\{i : \lambda_i(A) = \lambda_m(A)\}, \quad \text{and} \quad \bar{m} = \max\{i : \lambda_i(A) = \lambda_m(A)\}.$$

Thus

$$\lambda_{\hat{m}-1}(A) > \lambda_{\hat{m}}(A) = \cdots = \lambda_m(A) = \cdots = \lambda_{\bar{m}}(A) > \lambda_{\bar{m}+1}(A)$$

(appropriately understood when $\hat{m} = 1$ or $\bar{m} = n$). Corresponding to A, fix an orthogonal matrix U diagonalizing A, which is to say satisfying

$$U^T A U = \operatorname{Diag}(\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)), \tag{3.2}$$

and let U_m be the submatrix of U consisting of the columns indexed by $\hat{m}, \hat{m} + 1, \ldots, \bar{m}$.

Theorem 3.3 (Directional derivatives). *Any matrix H in $\mathbf{S}^n$ satisfies*

$$\lambda_m'(A; H) = \lambda_{m-\hat{m}+1}\big(U_m^T H U_m\big). \tag{3.4}$$

Furthermore, the sum of the eigenvalues coinciding with λ_m at A,

$$\sum_{i=\hat{m}}^{\bar{m}} \lambda_i,$$

is analytic near A.

Formula (3.4) follows easily from expressions for the directional derivative of the eigenvalue sum σ_m appearing in [11, 18]. The analyticity result appears in [21].

Observe that if $\hat{m} = m$ then the above result shows σ_{m-1} is analytic near A, and since σ_m is convex we deduce λ_m is Clarke regular at A. (In fact the converse is also true [14].) A similar argument shows $-\lambda_m$ is Clarke regular at A if $\bar{m} = m$ (and the converse is again true). In particular, $\lambda_{\hat{m}}$ and $-\lambda_{\bar{m}}$ are both Clarke regular at A.

4. Subdifferentials at zero

We begin our subdifferential calculations by computing directional derivatives of the mth largest eigenvalue λ_m at zero. We know these directional derivatives are sublinear approximants of λ_m, so we start with some cruder estimates.

Clearly we have the following sublinear majorants of λ_m, for all m:

$$\lambda_m \le \frac{\sigma_m}{m} \le \lambda_1.$$

Furthermore, for $m < n$ we have the concave minorants

$$\lambda_m \ge \frac{\sigma_n - \sigma_m}{n - m} \ge \lambda_n.$$

Proposition 4.1. *If $1 < m < n$, the function λ_m has no linear minorant or majorant.*

Proof: Suppose the matrix C in $\mathbf{S}^n$ satisfies

$$\langle C, G\rangle \le \lambda_m(G) \quad \text{for all } G \text{ in } \mathbf{S}^n.$$

Choose an orthogonal matrix U diagonalizing C:

$$U^T C U = \mathrm{Diag}(c_1, c_2, \ldots, c_n) \quad \text{for some vector } c \text{ in } \mathbf{R}^n.$$

Then by choosing $G = U(\mathrm{Diag}\, x)U^T$ in the above inequality (for vectors x in $\mathbf{R}^n$), we deduce

$$c^T x \le \lambda_m(\mathrm{Diag}\ x).$$

Now choosing x to be each standard basis vector in turn shows $c \leq 0$, but choosing $x = (-1, -1, \dots, -1)$ gives a contradiction. The argument for majorants is similar. $\square$

The above result shows the convex and concave subdifferentials

$$\partial\lambda_m(A) = \{C \in \mathbf{S}^n : \langle C, G - A\rangle \leq \lambda_m(G) - \lambda_m(A) \text{ for all } G \in \mathbf{S}^n\},$$
$$-\partial(-\lambda_m)(A) = \{C \in \mathbf{S}^n : \langle C, G - A\rangle \geq \lambda_m(G) - \lambda_m(A) \text{ for all } G \in \mathbf{S}^n\}$$

are both empty at $A = 0$, when $1 < m < n$. A similar argument shows the same result at an arbitrary matrix A in $\mathbf{S}^n$, and furthermore, that these subdifferentials remain empty at A for any function having the form $\lambda_m(\cdot) + o(\cdot - A)$ near A. Geometrically, the functions λ_m are very ‘bumpy’.

Recall (cf. (3.1)) the formula for the convex subdifferential of λ_1 at zero:

$$\partial\lambda_1(0) = \{C \in \mathbf{S}^n : C \text{ positive semidefinite, tr } C = 1\}.$$

Notice $\partial\lambda_1(0) = -\partial(-\lambda_n)(0)$.

Theorem 4.2 (Subdifferentials at zero). *The Clarke and Michel-Penot subdifferentials of the mth largest eigenvalue λ_m for $m = 1, 2, \dots, n$ all coincide at zero*:

$$\partial^{\diamond}\lambda_m(0) = \partial^{\mathrm{cl}}\lambda_m(0) = \partial\lambda_1(0).$$

Proof: The first equality follows from the positive homogeneity of λ_m (see formula (2.2)). To prove the result we, therefore, need to show any matrix G in $\mathbf{S}^n$ satisfies

$$\lambda_m^{\diamond}(0; G) = \lambda_1'(0; G),$$

or, using formula (2.2),

$$\sup_{H \in \mathbf{S}^n} \{\lambda_m(H + G) - \lambda_m(H)\} = \lambda_1(G). \tag{4.3}$$

On the one hand, the inequality

$$\lambda_m(H + G) \leq \lambda_m(H) + \lambda_1(G)$$

is a classical inequality of Weyl (see [12, Theorem 4.3.7], or [1]). To see the supremum is attained, choose an orthogonal matrix U with $U^T G U = \mathrm{Diag}\,\lambda(G)$ and any real $\alpha > \lambda_1(G) - \lambda_n(G)$. Then if we set

$$H = U \,\mathrm{Diag}\,(0, 0, \dots, 0, \underbrace{\alpha, \alpha, \dots, \alpha}_{m-1 \text{ terms}})\, U^T$$

we deduce

$$\lambda_m(H+G) - \lambda_m(H) = \lambda_m(U^T(H+G)U) - 0 = \lambda_m(\text{Diag}\, d),$$

where $d \in \mathbf{R}^n$ is the vector

$$(\lambda_1(G), \lambda_2(G), \ldots, \lambda_{n-m+1}(G), \lambda_{n-m+2}(G) + \alpha, \lambda_{n-m+3}(G) + \alpha, \ldots, \lambda_n(G) + \alpha).$$

The right-hand side of the equation is just the mth largest component of d, which is $\lambda_1(G)$, as required. □

5. Subdifferentials away from zero

The Michel-Penot and Clarke subdifferentials of the mth largest eigenvalue λ_m coincide at zero simply by positive homogeneity. Away from zero, their equality is no longer immediate. To prove these two subdifferentials are indeed equal we make the following steps for a fixed matrix A in $\mathbf{S}^n$.

- Use formula (3.4) to identify the usual directional derivative $\lambda_m'(A; \cdot)$.
- Calculate the best sublinear majorant of $\lambda_m'(A; \cdot)$ (which is exactly the Michel-Penot directional derivative $\lambda_m^\diamond(A; \cdot)$), using formula (2.3).
- Regularize $\lambda_m'(A; \cdot)$, using formula (2.1), to identify the Clarke directional derivative $\lambda_m^\circ(A; \cdot)$.

In general, $\lambda_m(A)$ belongs to a block of equal eigenvalues. The leading eigenvalue in this block is $\lambda_{\hat{m}}$, and its usual directional derivative at A is sublinear, by formula (3.4). We show its subdifferential coincides exactly with the Michel-Penot and Clarke subdifferentials of λ_m at A.

Theorem 5.1 (Michel-Penot Subdifferential). *The Michel-Penot subdifferentials of the eigenvalues λ_m at a matrix A in $\mathbf{S}^n$ coincide for all choices of m corresponding to equal eigenvalues*:

$$\partial^\diamond \lambda_m(A) = \partial(\lambda_{\hat{m}}'(A, \cdot))(0) \quad \textit{for all } m.$$

Proof: We wish to show that any matrix G in $\mathbf{S}^n$ satisfies

$$\lambda_m^\diamond(A; G) = \lambda_{\hat{m}}'(A; G)$$

(since $\lambda_{\hat{m}}'(A; \cdot)$ is sublinear). By formula (2.3), this is equivalent to

$$\sup_{H \in \mathbf{S}^n} \{\lambda_m'(A; H+G) - \lambda_m'(A; H)\} = \lambda_{\hat{m}}'(A; G).$$

Choose an orthogonal matrix U diagonalizing A as in Eq. (3.2). Then the directional derivative formula (3.4), applied in turn to λ_m and to $\lambda_{\hat{m}}$, shows the equation above is equivalent to

$$\sup_{H \in \mathbf{S}^n} \left\{ \lambda_{m-\hat{m}+1}\big(U_m^T (H+G) U_m\big) - \lambda_{m-\hat{m}+1}\big(U_m^T H U_m\big) \right\} = \lambda_1\big(U_m^T G U_m\big).$$

Since any matrix F in $\mathbf{S}^{\bar{m}-\hat{m}+1}$ satisfies

$$U_m^T \big(U_m F U_m^T\big) U_m = F,$$

this is equivalent to

$$\sup_{F \in \mathbf{S}^{\bar{m}-\hat{m}+1}} \left\{ \lambda_{m-\hat{m}+1}\big(F + U_m^T G U_m\big) - \lambda_{m-\hat{m}+1}(F) \right\} = \lambda_1\big(U_m^T G U_m\big),$$

and this follows from the proof of Theorem 4.2—see Eq. (4.3). □

It is not difficult to see

$$\partial(\lambda'_{\hat{m}}(A; \cdot))(0) = -\partial(-\lambda'_{\bar{m}}(A; \cdot))(0).$$

The next result gives a clearer interpretation of our calculation above. We use the notation conv to denote the convex hull of a set.

Proposition 5.2. *For any matrix A in $\mathbf{S}^n$, denote the eigenspace corresponding to the mth largest eigenvalue $\lambda_m(A)$ by $\mathbf{E}_m(A) \subset \mathbf{R}^n$. Then*

$$\partial(\lambda'_{\hat{m}}(A; \cdot))(0) = \operatorname{conv}\{xx^T : x \in \mathbf{E}_m(A),\ \|x\| = 1\}.$$

Proof: We wish to show that any matrix G in $\mathbf{S}^n$ satisfies

$$\lambda'_{\hat{m}}(A; G) = \max\{\langle G, xx^T \rangle : x \in \mathbf{E}_m(A),\ \|x\| = 1\}.$$

As in the proof of the previous result, we can write this

$$\lambda_1\big(U_m^T G U_m\big) = \max\{x^T G x : x \in \mathbf{E}_m(A),\ \|x\| = 1\}.$$

Since the columns of the matrix U_m are an orthonormal basis for $\mathbf{E}_m(A)$, the right-hand side is

$$\max\{(U_m y)^T G (U_m y) : y \in \mathbf{R}^{\hat{m}-\bar{m}+1},\ \|y\| = 1\},$$

which equals the left-hand side. □

We end this section with our main result.

Theorem 5.3 (Michel-Penot and Clarke Subdifferentials). *For any matrix A in $\mathbf{S}^n$ with eigenspace $\mathbf{E}_m(A) \subset \mathbf{R}^n$ corresponding to the mth largest eigenvalue $\lambda_m(A)$, the Michel-Penot and Clarke subdifferentials of λ_m at A are equal*:

$$\partial^{\diamond}\lambda_m(A) = \partial^{\mathrm{cl}}\lambda_m(A) = \mathrm{conv}\{xx^T : x \in \mathbf{E}_m(A),\ \|x\| = 1\}. \tag{5.4}$$

These subdifferentials all coincide for all choices of m corresponding to equal eigenvalues.

Proof: We know $\partial^{\diamond}\lambda_m(A) \subset \partial^{\mathrm{cl}}\lambda_m(A)$. For the opposite inclusion, fix a matrix G in $\mathbf{S}^n$ and an index m. By the regularization formula (2.1) there is a sequence of matrices A_k approaching A in $\mathbf{S}^n$ with

$$\lambda_m^{\circ}(A; G) = \lim_{k\to\infty} \lambda_m'(A_k; G).$$

By taking a subsequence we can assume the set of indices

$$\{i : \lambda_i(A_k) = \lambda_m(A_k)\}$$

is independent of the index k: we denote it by I. Since the eigenvalues λ_i are all continuous, observe

$$I \subset \{\hat{m}, \hat{m}+1, \ldots, \bar{m}\}. \tag{5.5}$$

For each k choose an orthogonal matrix U^k diagonalizing A_k: that is,

$$(U^k)^T A_k U^k = \mathrm{Diag}\,\lambda(A_k).$$

Again taking a subsequence we can assume U^k approaches an orthogonal matrix U which diagonalizes A, since λ is continuous: $U^T A U = \mathrm{Diag}\,\lambda(A)$.

Denote the submatrices of U and U^k with columns indexed by I by U_I and U_I^k. Now, by the Directional Derivatives Theorem (3.3),

$$\lambda_m'(A_k; G) \le \lambda_1\big((U_I^k)^T G U_I^k\big) \to \lambda_1\big(U_I^T G U_I\big).$$

If, as usual, we denote by U_m the submatrix of U with columns indexed by the right-hand side set in inclusion (5.5), then we see

$$\begin{aligned}\lambda_1\big(U_I^T G U_I\big) &= \max_{\|z\|=1} (U_I z)^T G (U_I z)\\ &\le \max_{\|y\|=1} (U_m y)^T G (U_m y)\\ &= \lambda_1\big(U_m^T G U_m\big)\\ &= \lambda_{\hat{m}}'(A; G),\end{aligned}$$

as in the proof of Theorem 5.1, whence

$$\lambda_m^\circ(A; G) \leq \lambda'_{\hat{m}}(A; G).$$

Since the matrix G was arbitrary, we deduce

$$\partial^{\text{cl}}\lambda_m(A) \subset \partial(\lambda'_{\hat{m}}(A; \cdot))(0),$$

and the result now follows by Theorem 5.1 and Proposition 5.2. (Alternatively, a more direct calculation also shows the opposite inclusion.) □

6. Comparison with earlier results

We can write the Clarke derivative formula (5.4) as

$$\lambda_m^\circ(A; G) = \max\{x^T G x : x \in \mathbf{E}_m(A),\ \|x\| = 1\},$$

a result first observed by Cox and Overton [6]. (A weaker inequality appeared in [9].) One proof of this result appears in [13], based on the almost-everywhere differentiability of λ_m. An alternative approach is developed in [14], this time via the Ioffe-Kruger-Mordukhovich ‘approximate subdifferential’. Our approach here is more elementary than either. The Michel-Penot result is, to our knowledge, new.

Cox [4] considered a continuously differentiable function $F : \mathbf{R}^k \to \mathbf{S}^n$, and studied the Clarke subdifferential of the composite function $\lambda_m \circ F$ (see also [2, 5] for related work). A rather complex argument leads to an exact formula when the eigenvalue of interest is the first or last in a block of equal eigenvalues, and just an inclusion otherwise: in our notation, for any index m and any points p and d in $\mathbf{R}^k$, setting $A = F(p)$, the result is

$$\begin{aligned}(\lambda_{\hat{m}} \circ F)^\circ(p; d) &= (\lambda_{\bar{m}} \circ F)^\circ(p; d) \\ &= \max\{x^T(F'(p)d)x : x \in \mathbf{E}_m(A),\ \|x\| = 1\} \\ &\leq (\lambda_m \circ F)^\circ(p; d).\end{aligned}$$

This now follows immediately from formula (5.4) by using the standard chain rule for the Clarke subdifferential (see [3]). Since, as we have observed, both $\lambda_{\hat{m}}$ and $-\lambda_{\bar{m}}$ are Clarke regular at A, the chain rule is exact for these eigenvalues, whereas in general it shows only an inclusion. The Michel-Penot subdifferential would permit a similar approach using a general differentiable function F.

One nice application of our main result, formula (5.4), is a transparent derivation of a well-known isotonicity property of eigenvalues (see [12, p. 181]): for matrices X and Y in $\mathbf{S}^n$,

$$X \succeq Y \Rightarrow \lambda(X) \geq \lambda(Y)$$

(where $X \succeq Y$ means $X - Y$ is positive semidefinite). To see this, note that for any index m the Lebourg Mean Value Theorem [3] implies the existence of a matrix Z in $\mathbf{S}^n$ such that

$$\begin{aligned}\lambda_m(X) - \lambda_m(Y) &\in \langle \partial\lambda_m(Z), X - Y\rangle \\ &= \operatorname{conv}\{\langle xx^T, X - Y\rangle : x \in \mathbf{E}_m(Z),\ \|x\| = 1\} \\ &= \operatorname{conv}\{x^T(X - Y)x : x \in \mathbf{E}_m(Z),\ \|x\| = 1\} \\ &\subset \mathbf{R}_+,\end{aligned}$$

whence the result.

References

1. R. Bhatia, "Perturbation bounds for matrix eigenvalues," Research Notes in Mathematics, vol. 162, Pitman, 1987.
2. T. Chatelain and M. Choulli, "Clarke generalized gradients for eigenvalues," Technical Report, University of Besançon, France, 1995.
3. F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley: New York, 1983.
4. S.J. Cox, "Extremal eigenvalue problems for the Laplacian," in Recent Advances in PDE, Proceedings, Summer School at El Escorial, 1992, Paris, M. Herrero and E. Zuazua (Eds.), 1993, Masson.
5. S.J. Cox, "The generalized gradient at a multiple eigenvalue," Journal of Functional Analysis, vol. 133, pp. 30–40, 1995.
6. S.J. Cox and M. Overton, 1994. Private communication.
7. H. Frankowska, "Inclusions adjointes associées aux trajectoires minimales d'inclusions différentielles," C.R. Acad. Sci. Paris, vol. 297, pp. 461–464, 1983.
8. E. Giner, "Ensembles et fonctions étoilés; applications à l'optimisation et au calcul différentiel généralisé," Technical Report, Paul Sabatier University, France, 1981.
9. B. Gollan, "Eigenvalue perturbation and nonlinear parametric optimization," Mathematical Programming Study, vol. 30, pp. 67–81, 1987.
10. J.-B. Hiriart-Urruty, "Miscellanies on nonsmooth analysis and optimization," in Nondifferentiable Optimization: Motivations and Applications, Workshop at Sopron, 1984, V.F. Demyanov and D. Pallaschke (Eds.), Lecture Notes in Economics and Mathematical Systems 255, Springer, 1985, pp. 8–24.
11. J.-B. Hiriart-Urruty and D. Ye,"Sensitivity analysis of all eigenvalues of a symmetric matrix," Numer. Math., vol. 70, pp. 45–72, 1995.
12. R.A. Horn and C. Johnson, Matrix Analysis, Cambridge University Press: Cambridge, UK, 1985.
13. A.S. Lewis, "Derivatives of spectral functions," Mathematics of Operations Research, vol. 6, pp. 576–588, 1996.
14. A.S. Lewis, "Nonsmooth analysis of eigenvalues," Mathematical Programming, 1998, to appear.
15. A.S. Lewis and M.L. Overton, "Eigenvalue optimization," Acta Numerica, vol. 5, pp. 149–190, 1996.
16. P. Michel and J.-P. Penot, "Calcul sous-différentiel pour les fonctions lipschitziennes et non lipschitziennes," C.R. Acad. Sci. Paris, vol. 298, pp. 269–272, 1984.
17. P. Michel and J.-P. Penot, "A generalized derivative for calm and stable functions," Differential and Integral Equations, vol. 5, pp. 433–454, 1992.
18. M.L. Overton and R.S. Womersley, "Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices," Mathematical Programming, Series B, vol. 62, pp. 321–357, 1993.
19. J.-P. Penot, "Variations on the theme of nonsmooth analysis: Another subdifferential," in Non-Differentiable Optimization: Motivations and Applications, Workshop at Sopron, 1984, V.F. Demyanov and D. Pallaschke (Eds.), Lecture Notes in Economics and Mathematical Systems 255, Springer, 1985, pp. 8–24.
20. E. Polak, "On the mathematical foundations of non-differentiable optimization in engineering design," SIAM Review, vol. 29, pp. 21–89, 1987.
21. N.-K. Tsing, M.K.H. Fan, and E.I. Verriest, "On analyticity of functions involving eigenvalues," Linear Algebra and its Applications, vol. 207, pp. 159–180, 1994.
22. L. Vandenberghe and S. Boyd, "Semidefinite programming," SIAM Review, vol. 38, pp. 49–95, 1996.