SOME REMARKS ON INVERSE AND EXTREMAL EIGENVALUE PROBLEMS

Michael L. Overton [1]

This paper gives a broad outline of recent work on both theoretical and numerical aspects of inverse and extremal eigenvalue problems, especially the developments that have taken place during my 1986-87 visit at the Centre for Mathematical Analysis and Mathematical Sciences Research Institute of the Australian National University. I would like to start by thanking all my hosts at the A.N.U. for their warm hospitality, especially Dr. Mike Osborne.

Let $A(x)$ be a smooth real $n \times n$ matrix function of a parameter vector $x \in \mathbb{R}^m$, and let $\lambda_i(x)$, $i = 1,...,n$, be the eigenvalues of $A(x)$. The inverse eigenvalue problem is :

IEP: Given $\lambda_i^* \in \mathbb{C}$, $i = 1,...,n$, and a domain $D \subseteq \mathbb{R}^m$, find $x \in D$ such that the sets $\{\lambda_i(x),\ i = 1,...,n\}$ and $\{\lambda_i^*,\ i = 1,...,n\}$, are the same. (We have not written $\lambda_i(x) = \lambda_i^*$ to avoid difficulties with inconsistent orderings.)

The most common version of the extremal eigenvalue problem is :

EEP: Given a domain $D \subseteq \mathbb{R}^m$, find $x \in D$ so that the spectral radius

$$\rho(x) = \max_{1 \le i \le n} |\lambda_i(x)|$$

is minimized over D.

The relationship between IEP and EEP is thus similar to the usual relationship between solving nonlinear systems of equations and nonlinear optimization. In particular, IEP may have no solution, while EEP must have a solution if D is compact, since $\rho(x)$ is continuous. The feature of IEP and EEP

---

[1]Visiting A.N.U., 1986–87. On leave from Courant Institute of Mathematical Sciences, New York University. This work was supported in part by the National Science Foundation under Grant DCR–85–02014.

which makes them much harder to solve than many systems of nonlinear equations and optimization problems is that the eigenvalues $\{\lambda_i(x)\}$ are not differentiable at points where they coalesce.

Let us first discuss IEP in the symmetric case, i.e., $A(x) = A(x)^T$ for all x. Because the eigenvalues are real, we may use the ordering $\lambda_1(x) \geq \ldots \geq \lambda_n(x)$ and also $\lambda_1^* \geq \ldots \geq \lambda_n^*$. Assume first that $\{\lambda_i^*\}$ are distinct, and that IEP has a solution x^*. Because the eigenvalues $\lambda_i(x)$ are each differentiable near x^*, it makes sense to assume $n = m$. Indeed, this is natural in many applications; for example, often $A(x)$ has fixed off-diagonal elements and variable diagonal elements $\{x_k\}$. Because the eigenvalues are differentiable, Newton's method is applicable. There are at least three possible systems of nonlinear equations to which Newton's method might be applied; let us write these as $f(x) = 0$, $g(x) = 0$ and $h(x) = 0$. The first and most natural is defined by

$$f_i(x) = \lambda_i(x) - \lambda_i^*, \; i = 1,\ldots,n. \tag{1}$$

This formulation has been used by many people; the earliest reference may be [1]. The second formulation is

$$g_i(x) = \det(A(x) - \lambda_i^* I), \; i = 1,\ldots,n; \tag{2}$$

see [2]. The third, due to Osborne [3], is more complicated to explain but may be the most computationally attractive in many situations. The reason for this is that obtaining the Jacobian of $f(x)$ and $g(x)$ requires computing the complete eigensystem of $A(x)$, which the third method avoids. Let $\{w_i\}$ be a set of n fixed independent vectors, which are supposed to be the best approximation to the eigenvectors of $A(x^*)$ known, say, at the ν^{th} iteration. Define $h(x)$ by

$$(A(x) - \lambda_i^* I)v_i(x) = h_i(x)w_i, \; i = 1,\ldots,n, \tag{3}$$

and

$$v_i(x)^T w_i = 1, \quad i = 1,\ldots,n. \tag{4}$$

Here $\frac{1}{h_i(x)}\, v_i(x)$ is the vector obtained by applying a single step of inverse iteration for each i, using the required eigenvalues $\{\lambda_i^*\}$ to shift $A(x)$ and the vectors $\{w_i\}$ for the right-hand sides. The quantity $h_i(x)$ is the scalar factor required to normalize $v_i(x)$ according to (4); $h_i(x)$ is in fact an approximation to (1). Note that since $\{w_i\}$ depends on the iteration count ν, the function $h(x)$ is differently defined at each step of the iteration, but this is not an unusual situation in the formulation of Newton's method. It turns out (see [3], [4], [5]) that differentiating (3), (4), gives a Jacobian of $h(x)$ which is closely related to the Jacobian of $f(x)$, the difference being that the role of the eigenvectors of $A(x)$ is replaced by the eigenvector approximations $\{v_i\}$ obtained from the inverse iteration steps. It is natural to set $w_i^{(\nu+1)} = v_i(x^{(\nu)})$, to define $h(x)$ at the beginning of the next iteration.

Ref. [5] gives a comprehensive study of methods for solving IEP, emphasizing four in particular. The ones known as Methods I and IV are simply the aforementioned Newton methods applied to (1) and (2). The one known as Method II is slightly different from Osborne's method just described. One difference lies in the choice of right-hand side of the linear system to be solved at each iteration (the left-hand side having the same Jacobian coefficient matrix). The other difference is the choice of normalization condition following the inverse iteration steps. Ref. [5] uses

$$v_i(x)^T v_i(x) = 1\ , \quad i = 1,\ldots,n, \tag{5}$$

instead of (4). Although in many ways this is more natural, the resulting method is not actually Newton's method applied to (3), (5), although it is quadratically convergent.

The other method, Method III of [5], was an original contribution. In some ways it is the most interesting, since it leads directly to the correct formulation of the methods when multiple eigenvalues are present. However, we shall not discuss it since it is described at some length in [6] as well as [5]. It is not known to be Newton's method applied to any particular function, although again it is quadratically convergent.

Practical experience shows that all of Methods I, II, III are superior to Method IV.

The main contribution of [5] was to show how to properly formulate IEP when the set $\{\lambda_i^*\}$, i = 1,...,n, contains multiple entries. In this case (1) is not differentiable, (2) is not a distinct set of equations and (3) is not well defined. Suppose for simplicity that only one multiple eigenvalue is imposed, say $\lambda_1^* = \ldots = \lambda_t^*$. The most important observation to make is that the resulting restricting set of equations on the parameter space is generically of dimension, not t, but $t(t+1)/2$. This fact goes back to von Neumann and Wigner in 1929 [7] and, in the context of requiring a symmetric matrix to have a given rank, Ledermann in 1937 [8], although it is not widely known. Consequently, to be well posed, a generic IEP problem must have the number of parameters, m, chosen to reflect this fact, i.e. $m = t(t+1)/2 + n - t$ if all eigenvalues are prescribed.

It turns out [5] that the appropriate generalization of (1) to which Newton's method should be applied is

$$f_{ij}^{(1)}(x) = q_i(x)^T A(x) q_j(x) - \delta_{ij}\lambda_i^*, \quad 1 \le i \le j \le t \tag{6}$$

$$f_i^{(2)}(x) = q_i(x)^T A(x) q_i(x) - \lambda_i^*, \quad t+1 \le i \le n. \tag{7}$$

Here $f(x)$ consists of two parts, both to be set to zero, with $t(t+1)/2$ and $n - t$ components respectively. The vectors $\{q_i(x)\}$ are an orthonormal basis of eigenvectors of $A(x)$. Note that (1) and (7) are actually the same, $i = t+1,\ldots,n$. One must be careful about what exactly is meant by differentiating (6); see [5].

We still do not know how to generalize formulation (2) to allow multiple eigenvalues. The key may be in the Ledermann arguement [8], which is as follows. We require $M = A(x) - \lambda_i^* I$ to have rank $n - t$. This is true if there is a principal submatrix M_1 of M with dimension $n - t$ and nonzero determinant, and if each principal submatrix of M with dimension $n - t + 1$ or $n - t + 2$ and leading block M_1 has zero determinant [9, p.79]. The last requirement consists of $t + \binom{t}{2}$ conditions, which is the correct number. The difficulty in exploiting this fact to apply Newton's method is that one does not generally know a suitable submatrix M_1. In

any case generalizing formulation (2) is only of academic interest since, when the eigenvalues are distinct, Method IV is inferior to the others in practice.

Ref. [5] also generalized Method II to allow multiple eigenvalues. Here we comment on a possible corresponding generalization of formulation (3). Let W_1 be an $n \times t$ matrix whose columns are t independent vectors approximating eigenvectors of $A(x^*)$ corresponding to $\lambda_1^* = \ldots = \lambda_t^*$. Define $H_1(x)$ by

$$(A(x) - \lambda_1^* I)V_1(x) = W_1 H_1(x) \tag{8}$$

where

$$V_1(x)^T V_1(x) = I \tag{9}$$

and $H_1(x)$ is $t \times t$ upper triangular. The idea is that $V_1(x)H_1(x)^{-1}$ is a QR factorization of $(A(x) - \lambda_1^* I)^{-1}W_1$. A QR factorization of this matrix was also used in the generalization described in [5]. The point we are making here is that it is possible to consider applying Newton's method directly to the inverse triangular factor $H_1(x)$, noting that it has the right number of equations to be set to zero, namely $t(t+1)/2$. However, the normalization equation (9) is a generalization of (5), not (4), and hence the relevant Jacobian does not nicely simplify as in Osborne's method [3]. A generalization of (4) would be

$$V_1(x)^T W_1 = I$$

which would seem to be an invalid normalization condition, giving the wrong number of equations.

So far we have discussed only symmetric inverse eigenvalue problems. These methods generalize without too much difficulty to the nonsymmetric case provided $A(x)$ is <u>not</u> <u>defective</u> at a solution x^*. We shall not pursue this further here, deferring the discussion of nonsymmetric problems to the end of the paper where we comment on the nonsymmetric EEP.

The Newton methods just described for IEP are useful as they stand only in a neighbourhood of a solution. Finding such a neighbourhood, or ascertaining whether a solution exists at all, are always difficult problems in the general solution of nonlinear

equations. In this sense minimization problems are generally easier to solve, since one can find a "lower" point and in that sense make progress towards finding at least a local minimum. Let us therefore turn to the minimization problem EEP. Most of my time at A.N.U. has been spent working on EEP, both the symmetric and nonsymmetric versions.

Let us first consider the symmetric version of EEP, i.e., $A(x) = A(x)^T$ for all x. Let us also assume that $A(x)$ is an affine function, which it is in many applications, to avoid the complication of minima introduced by nonlinearities in $A(x)$. Then $\rho(x)$ is convex, since the largest eigenvalue of a symmetric matrix is a convex function of the matrix elements [10]. However, $\rho(x)$ is not differentiable at points x where the maximum eigenvalue, $\lambda_1(x)$, is multiple. In fact, $\rho(x)$ is not differentiable even if $\lambda_1(x)$ and $\lambda_n(x)$ are both distinct with a common modulus, but this sort of nondifferentiability, where $\rho(x)$ is simply a pointwise maximum of two differentiable functions, is easily analyzed [11]. Let us suppose that

$$\rho(x) = \lambda_1(x) = \ldots = \lambda_t(x) > \lambda_{t+1}(x) \geq \ldots \geq \lambda_{n-s}(x) > \lambda_{n-s+1}(x) = \ldots = \lambda_n(x) = -\rho(x)$$

so that t and s are the multiplicities of the two eigenvalues achieving the maximum modulus. If t or s is greater than one, determining whether x is a minimizer of f is not an easy task. This contrasts with IEP where, given x, determining whether x solves IEP is trivial. Since $\rho(x)$ is convex, Rockafellar's theory of subgradients [12] applies. Even with the help of this powerful theory, however, obtaining a computationally verifiable optimality condition is not trivial. In [13] we have given such a verifiable optimality condition, closely following the work of Fletcher on a related topic [10]. Fletcher's work in turn is based on the Rockafellar theory. The optimality condition is as follows :

THEOREM *Assume that* $A(x)$ *is an affine symmetric matrix function. A necessary and sufficient condition for* x *to minimize* $\rho(x)$ *over* $D = \mathbb{R}^m$ *is that there exist "dual matrices"* U *and* V, *of dimension* $t \times t$ *and* $s \times s$ *respectively, with* U *and* V *both symmetric and positive semi-definite and satisfying*

$$\text{tr } U + \text{tr } V = 1 \tag{10}$$

$$Q_1^T \frac{\partial A(x)}{\partial x_k} Q_1 : U - Q_2^T \frac{\partial A(x)}{\partial x_k} Q_2 : V = 0 \ , \ k = 1,\ldots,m \tag{11}$$

where Q_1 and Q_2 are respectively $n \times t$ and $n \times s$ matrices with orthonormal columns which are eigenvectors for the multiple eigenvalues $\lambda_1(x) = \ldots = \lambda_t(x)$ and $\lambda_{n-s+1}(x) = \ldots = \lambda_n(x)$ respectively. The ":" is the matrix inner product operator defined by

$$A:B = \text{tr}(AB)$$

where A and B are symmetric matrices of the same dimension. □

The interesting thing about this theorem is that it involves "dual matrices". These generalize the notion of dual variables (Lagrange multipliers) familiar from linear and nonlinear programming. In such problems the dual variables are generally required to be nonnegative for optimality. Here we have the generalized condition that the dual matrices must be nonnegative definite, i.e., have nonnegative eigenvalues. The procedure for resolving optimality is simply to solve (10), (11) (see below) for $U = U^T$, $V = V^T$, and then check whether U and V are nonnegative definite.

The main theoretical contribution of [13] was to show further how, if the optimality condition is not satisfied at x, a descent direction may be generated by splitting one of the multiple eigenvalues. It is the indefinite dual matrix that gives the key to finding a descent direction, just as it is a negative dual variable ("reduced cost") which gives the information required to obtain a descent direction from a vertex in linear programming. See [13] for details.

One important feature of the symmetric EEP is that it is common for problems to have a solution where t or s (or both) is greater than one. The reason for this is that the goal of EEP is to find a point x such that the eigenvalues of $A(x)$ are squeezed into as small an interval $[-\rho,\rho]$ as possible. Typically, as the optimization proceeds, more eigenvalues coalesce to the values $\pm\rho$, increasing t and s

accordingly. However, the von Neumann-Wigner argument imposes a generic limit on t and s, namely

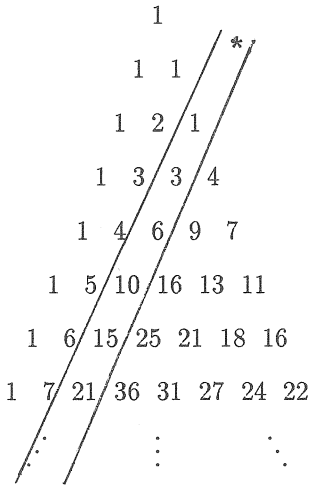
$$\frac{t(t+1)}{2} + \frac{s(s+1)}{2} \leq m + 1 \tag{12}$$

where the "1" on the right-hand side reflects the fact that the common modulus ρ, like the m parameters $\{x_k\}$, is free. Problems for which (12) is violated are said to be degenerate. Note that since the linear system (10), (11) is $m + 1$ equations in $t(t+1)/2 + s(s+1)/2$ unknowns, it is generically solvable if (12) holds with equality. If (12) holds with inequality and (10), (11) is not a solvable system, it is generically the case that a descent direction may be found without splitting a multiple eigenvalue.

It is also possible that $A(x)$ has no multiple eigenvalue at the solution. This happens, for example, if $A(x)$ is tridiagonal with fixed nonzero off-diagonal elements.

Along with the theoretical contributions just described, [13] gives a practical algorithm for minimizing $\rho(x)$ in the symmetric case. The method is based on successive quadratic programming (QP), using estimates of the multiplicities t and s to incorporate the appropriate linear constraints in the QP's. These linear constraints are effectively the same linearizations used in the Newton method applied to (6), (7) in the context of IEP. If (12) holds with equality, the minimum is well-defined by first-order information and it is sufficient to use successive linear programming instead of successive QP. Otherwise, second-order information must be used in the QP's to give a method which is locally quadratically convergent. For details and numerical examples, see [13].

Clearly, (12) is not solvable with equality for all m. Jerry Kautsky of Flinders University has pointed out that a nice way to display the integers m for which (12) is solvable with equality is the "modified Pascal triangle":



Here each number in the diagonal * has the form $t(t+1)/2$, and each number below the diagonal * is obtained by adding two numbers in the diagonal * . The number of such possible integers expressed as a percentage of all integers from 1 to $\hat{m}$, say, decreases to zero as $\hat{m} \to \infty$.

Finally, let us comment on the nonsymmetric EEP. In this case, $\rho(x)$ is not convex. Furthermore, $\rho(x)$ is not differentiable or even Lipschitz at points where $A(x)$ has multiple eigenvalues achieving the maximum modulus. Since $A(x)$ is a real matrix, its eigenvalues are either real or occur in complex conjugate pairs, and hence it will have (at least one) double real eigenvalue at points x where a pair of eigenvalues cross from being real to being a complex conjugate pair. It is therefore important, as in the symmetric case, to be able to resolve the question of optimality, and to be able to find descent directions when they exist, at points of multiple eigenvalues. Whether $\rho(x)$ tends to be minimized at points of multiple eigenvalues (as in the symmetric case) is not so clear. Many types of minima are possible, for example, ordinary "max function" minima where several distinct eigenvalues (either real or complex conjugate pairs) all achieve the same maximum modulus. In such cases optimality can be resolved by standard min-max techniques [11]. If we suppose that $D = \mathbb{R}^m$ and that only one distinct eigenvalue achieves the maximum modulus, my impression is that minima are most likely to occur at points where $A(x)$ has a nonderogatory

multiple eigenvalue, that is one where the corresponding part of the Jordan form of $A(x)$ is a single Jordan block. It seems to be very hard to resolve optimality, at such points, or indeed any points where a maximum eigenvalue (in modulus) has any non-trivial Jordan blocks, i.e. the eigenvalue is defective.

However, my recent work with Rob Womersley of U.N.S.W. [14] does show how to resolve optimality and, if not optimal, find a descent direction, at points x where $A(x)$ has (possibly many, real or complex) multiple but nondefective eigenvalues achieving the maximum modulus. Even in this case $\rho(x)$ is not locally Lipschitz, and neither the theory of subgradients nor Clarke's theory of generalized gradients [15] is applicable. We give a necessary and sufficient optimality condition which involves dual matrices, as in the symmetric case. In the case of one multiple nondefective eigenvalue achieving the maximum modulus, the optimality condition is that the associated dual matrix must be a multiple of the identity matrix; otherwise, a descent direction may be found. See [14] for details.

There are still many difficulties to be resolved before an algorithm for the nonsymmetric EEP becomes a reality. Not only does the theoretical question of optimality remain open in the defective case, but there are many practical and numerical difficulties. For example, even computing the Jordan form of $A(x)$ at a single point x is well known to be a hard problem numerically. Also, minimizing $\rho(x)$ may not be a reasonable goal, since if $A(x)$ is defective at a solution x, a perturbation in x of size ε may increase the spectral radius $\rho(x)$ by $O(\varepsilon^{1/t})$, where t is the order of the largest Jordan block in the Jordan form of $A(x)$. Thus Jerry Kautsky has suggested that one may really prefer to minimize say $\tilde{\rho}(x)$, where

$$\tilde{\rho}(x) = \sup_{\|d\| \le \varepsilon} \rho(x+d)$$

and ε is a given tolerance. This problem may be even harder to solve than EEP.

Let us briefly summarize. In the symmetric case, methods are well understood for IEP, even in the multiple eigenvalue case, but they are local methods and require a reasonably good starting point. Optimality conditions are well understood for EEP,

and a practical algorithm is now available [13]. In the nonsymmetric case, defective eigenvalues impose great difficulties. However, new necessary and sufficient optimality conditions are now known for EEP in the nondefective case [14].

REFERENCES

[1] A.C. Downing and A.S. Householder, "Some inverse characteristic value problems, J. Assoc. Comp. Mach. 3 (1956) 203-207.

[2] F.W. Biegler-König, "A Newton iteration process for inverse eigenvalue problems", Numer. Math. 37 (1981) 349-354.

[3] M.R. Osborne, "On the inverse eigenvalue problem for matrices and related problems for difference and differential equations", Conference on Application of Numerical Analysis, Lecture Notes in Mathematics No.228, 1971, Springer-Verlag.

[4] O. Hald, "On discrete and numerical Sturm-Liouville problems", Ph.D. Dissertation, Dept. of Math., New York University, 1972.

[5] S. Friedland, J. Nocedal and M.L. Overton, "The formulation and analysis of numerical methods for inverse eigenvalue problems", SIAM J. Numer. Anal., 24 (1987) 634-667.

[6] S. Friedland, J. Nocedal and M.L. Overton, "Four quadratically convergent methods for solving inverse eigenvalue problems", in : D.F. Griffiths and G.A. Watson, eds., Numerical Analysis (Proceedings of the Eleventh Dundee Biennial Conference on Numerical Analysis, June 1985), Pitman Research Notes in Math. 140, Wiley, New York, 1986, pp.47-65.

[7] J. von Neumann and E. Wigner, "Über das Verhalten von Eigenwerten bei adiabatischen Prozessen", Physik Zeitschr. 30 (1929) 467-470.

[8] W. Ledermann, "On the rank of the reduced correlational matrix in multiple-factor analysis", Psych. 2 (1937) 85-93.

[9] L.E. Dickson, "Modern Algebraic Theories", Sanborn, Chicago (1926).

[10] R. Fletcher, "Semi-definite matrix constraints in optimization", SIAM J. on Control and Optimization 23 (1985) 493-513.

[11] R. Fletcher, "Practical Methods of Optimization", Vol.2, Wiley, Chichester and New York.

[12] R.T. Rockafellar, "Convex Analysis", Princeton University Press, Princeton, N.J., 1970.

[13] M.L. Overton, "On minimizing the maximum eigenvalue of a symmetric matrix", Centre for Mathematical Analysis Report CMA-R03-87, Australian National University, Canberra (1987). To appear in : Linear Algebra in Signals, Systems and Control (B.N. Datta, Ed.). Soc. Industr. Appl. Math., Philadelphia, (1988).

[14] M.L. Overton and R.S. Womersley, "On minimizing the spectral radius of a nonsymmetric matrix function - optimality conditions and duality theory", Centre for Mathematical Analysis Report, CMA-R26-87, Australian National University, Canberra, (1987). To appear in SIAM J. Matrix Anal. Applic. (1988).

[15] F.H. Clarke, "Optimization and Nonsmooth Analysis", Wiley-Interscience, New York, 1983.

Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York
N Y 10012
U.S.A.