Optimizing Eigenvalues of Symmetric Definite Pencils[1]

Jean-Pierre A. Haeberly
Department of Mathematics
Fordham University
Bronx, NY 10458
718–817–3224 haeberly@murray.fordham.edu

Michael L. Overton
Computer Science Department
Courant Institute of Mathematical Sciences
New York University
251 Mercer St., New York, NY 10012
212–998–3121 overton@cs.nyu.edu

Abstract

We consider the following quasiconvex optimization problem: minimize the largest eigenvalue of a symmetric definite matrix pencil depending on parameters. A new form of optimality conditions is given, emphasizing a complementarity condition on primal and dual matrices. Newton's method is then applied to these conditions to give a new quadratically convergent interior-point method which works well in practice. The algorithm is closely related to primal-dual interior-point methods for semidefinite programming.

1. Introduction

Many matrix inequality problems in control can be cast in the form: minimize the maximum eigenvalue of the Hermitian definite pencil $(A(x), B(x))$, w.r.t. a parameter vector x, subject to positive definite constraints on $B(x)$ and sometimes also on other Hermitian matrix functions of x. The maximum eigenvalue is a quasiconvex function of the pencil elements and therefore of the parameter vector x if A, B depend affinely on x. This quasiconvexity reduces to convexity in the important special case that $B(x) \equiv I$, i.e. the pencil reduces to an affine matrix function $A(x)$. In this case, the eigenvalue optimization problem is equivalent to semidefinite programming (SDP), i.e. the problem of minimizing a linear function subject to semidefinite constraints on linear matrix families. A duality theory, completely analogous to the theory of linear programming (LP) is known for SDP [1]. In the general case as well as in the special case of SDP, the eigenvalues are generally not differentiable at the solution point. This is because the eigenvalues of a matrix or pencil fail to be differentiable functions of its elements when multiple eigenvalues are present, which is normally the case at the optimal solution of practical problems.

We discuss optimality conditions and algorithms, and our practical experience with these algorithms. For simplicity we restrict the discussion to the case of real symmetric pencils, although extension to the Hermitian case is not difficult.

2. Optimality conditions

Let $A \bullet B = \mathbf{tr}\, AB$ denote the Frobenius inner product of two symmetric matrices. By $A \succeq 0$ we mean A is positive semidefinite. Assume that $A(x)$ and $B(x)$ are continuously differentiable functions of $x \in \Re^m$, taking real symmetric n by n matrix values, with $B(x)$ positive definite for all x. Let $A_k(x)$, $B_k(x)$ denote the partial derivatives of $A(x)$, $B(x)$ with respect to x_k. In what follows we shall suppress the dependence of various quantities on the parameter vector x. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of the symmetric pencil $(A, B) = (A(x), B(x))$, i.e. solutions of $\det(\lambda B - A) = 0$. Let Q be an n by n matrix of eigenvectors, i.e. satisfying $AQ = \mathrm{Diag}(\lambda_i)BQ$, with the normalization condition $Q^T BQ = I$.

[1] Computer Science Dept. Report 660, Courant Institute of Mathematical Sciences, New York University, March 1994. Available by anonymous ftp to cs.nyu.edu in the file pub/local/overton/acc94.ps. To appear in the Proceedings of the 1994 American Control Conference, Baltimore, June 30-July 2, 1994.

The following result is given in [2].

Theorem 1. Assume that the multiplicity of $\lambda_1 = \lambda_1(x)$ is known to be t, and let Q_1 be the n by t matrix whose columns are the corresponding t columns of the eigenvector matrix Q. Then a necessary condition for x to minimize λ_1 is that there exists a positive semidefinite symmetric t by t matrix V, such that

$$\begin{aligned} \mathbf{tr}\, V &= 1 && (1)\\ V \bullet (Q_1^T E_k Q_1) &= 0, \quad k = 1, \ldots, m. && (2) \end{aligned}$$

where $E_k = \lambda_1 B_k - A_k$.

The result in Theorem 1 requires the knowledge of the optimal multiplicity t. In the following result, one does not need to know the multiplicity t.

Theorem 2. A necessary condition for x to minimize λ_1 is that there exists a positive semidefinite symmetric n by n matrix U satisfying

$$\begin{aligned} \mathbf{tr}\, U &= 1, && (3)\\ U \bullet (HE_kH^T) &= 0, \quad k = 1, \ldots, m, && (4)\\ U(HEH^T) &= 0, && (5) \end{aligned}$$

where

$$\begin{aligned} E_k &= \lambda_1 B_k - A_k, && (6)\\ E &= \lambda_1 B - A, && (7) \end{aligned}$$

and H is any matrix satisfying $HBH^T = I$, e.g. the inverse symmetric square root or the inverse Cholesky factor of B.

Proof. Let Q be defined as above. We have $Q^T B Q = I$, $Q^T A Q = \mathrm{Diag}(\lambda_i)$ and $Q^T E Q = \mathrm{Diag}(\eta_i)$, where $\eta_i = \lambda_1 - \lambda_i \geq 0$. Define $P = H^{-T} Q$, which is an orthogonal matrix. Let t be the multiplicity of λ_1, so $\eta_1 = \cdots = \eta_t = 0$, and let D be the diagonal $(n-t)$ by $(n-t)$ matrix $\mathrm{Diag}(\eta_{t+1}, \ldots, \eta_n)$. Let V denote a symmetric t by t positive semidefinite matrix satisfying the conclusions of Theorem 1, and define $\tilde{V}$ to be the n by n matrix

$$\tilde{V} = \begin{bmatrix} V & 0 \\ 0 & 0 \end{bmatrix}.$$

Finally, write $U = P\tilde{V}P^T$. Then $\mathbf{tr}\, U = \mathbf{tr}\, V = 1$. Moreover, it is easily verified that for any $1 \leq k \leq m$

$$U \bullet (HE_kH^T) = V \bullet (Q_1^T E_k Q_1) = 0.$$

Finally, we verify that U satisfies (5). Observe that

$$U(HEH^T) = 0 \iff P^T(U(HEH^T))P = 0.$$

But

$$\begin{aligned} P^T(U(HEH^T))P &= \tilde{V}(Q^T(\lambda_1 B - A)Q)\\ &= \begin{bmatrix} V & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix}\\ &= 0. \end{aligned}$$

A key component of Theorem 2 is the *complementarity condition* (5). Let us consider the matrices U and HEH^T further. The eigenvalues of HEH^T are $0 = \eta_1 \leq \cdots \leq \eta_n$, i.e. the eigenvalues of the shifted pencil (E, B). Let $\theta_1 \geq \cdots \geq \theta_n \geq 0$ denote the eigenvalues of U. We see that from the definition of U, $\theta_{t+1} = \cdots = \theta_n = 0$. Thus

$$\eta_i \theta_i = 0, \quad i = 1, \ldots, n. \qquad (8)$$

Furthermore, let W be a t by t orthogonal matrix which diagonalizes V, i.e. $W^T V W = \mathrm{Diag}(\theta_1, \ldots \theta_t)$. Then

$$P \begin{bmatrix} W & 0 \\ 0 & I \end{bmatrix}$$

is an eigenvector matrix for both HEH^T and U, i.e. it diagonalizes both matrices simultaneously. We refer to HEH^T as the *primal* matrix and U as the *dual* matrix. Recall that two symmetric matrices commute if and only if they share a basis of eigenvectors. It follows that the complementarity condition (5) holds if and only if the primal and dual matrix commute and their eigenvalues satisfy (8).

In the special case of SDP, i.e. if $A(x)$ is affine and $B(x) \equiv I$, Theorem 2 is known in the context of the duality theory given in [1]. In the general case, there is no such general duality theory, but it is remarkable that the complementarity form of the optimality conditions still hold as expressed in Theorem 2.

3. Interior Point Methods

In recent years it has been realized that the interior point methods which have been so successful for LP can be extended to solve eigenvalue optimization problems. For optimizing the eigenvalues of pencils, see [3,4,5]. These methods are based on Huard's method of centers and they consist of an "outer iteration", each step of which requires the solution of a nonlinear problem using an "inner iteration". Usually Newton's method is used to solve this nonlinear problem, in which case each step of the inner iteration requires factoring a dense Hessian matrix whose order is m, the number of unknowns. The method of [6] is a modified method of centers for which the objective values of the outer iteration converge quadratically. Note that this method still requires a nonlinear problem to be solved by an inner iteration at every step of the outer iteration.

In the case that $B(x) = I$ and $A(x)$ is affine, the eigenvalue optimization problem reduces to SDP and the options for interior point methods become more

numerous. In [1] it is argued that essentially any interior point method designed for LP can be extended to solve SDP. In LP it is now generally agreed that primal-dual interior-point methods are especially efficient [7]. A specific primal-dual method for SDP with a proof of global convergence was given by [8]. A related method was given by [9]. A different approach to primal-dual interior point methods for SDP is given in [10].

4. Quadratically Convergent Local Methods

In [2], the authors derived a quadratically convergent local method for optimizing eigenvalues of pencils. This method extended earlier work on optimizing eigenvalues of matrices [11,12]. Note that each step of this algorithm requires only the solution of a linear system of equations, though the form of the equations is quite complicated. Even in the case of matrix eigenvalue optimization, the proof of quadratic convergence is nontrivial [13], since the method cannot be described as the straightforward application of Newton's method to a nonlinear system. The method verifies optimality by explicitly computing the matrix V given in Theorem 1. The primary limitation of this method is that it is necessary to guess the optimal eigenvalue multiplicity t before it can be applied. A secondary limitation is that $B(x)$ must be affine, though $A(x)$ need not be.

We now give a new quadratically convergent local method for optimizing eigenvalues of pencils, based on Theorem 2. Because there is no need to guess the eigenvalue multiplicity, this method seems much more practical than that of [2]. The new method can be viewed as a local primal-dual interior-point method. This is because primal-dual interior-point methods for LP and SDP can be viewed as applying Newton's method to a set of nonlinear equations which define primal feasibility, dual feasibility, and complementarity [7,9,10].

Let **vec** denote the map from the space $S\Re^{n\times n}$ of symmetric matrices onto $\Re^{n(n+1)/2}$ satisfying

$$M \bullet N = \mathbf{vec}(M)^T \mathbf{vec}(N)$$

for any $M, N \in S\Re^{n\times n}$. Let $l = n(n+1) + m + 1$ and let $z = (x, \omega, h, u) \in \Re^l$ where $x \in \Re^m$, $\omega \in \Re$, and $h, u \in \Re^{n(n+1)/2}$. The optimality conditions of Theorem 2 can be written as

$$F(x, \omega, h, u) = 0, \quad \omega B(x) - A(x) \succeq 0, \quad U \succeq 0$$

where $F : \Re^l \longrightarrow \Re^l$ is the nonlinear map with $F(z) =$

$$\begin{pmatrix} U \bullet H(\omega B_1(x) - A_1(x))H \\ \vdots \\ U \bullet H(\omega B_m(x) - A_m(x))H \\ 1 - \mathbf{tr}\, U \\ \mathbf{vec}(I - HB(x)H) \\ \mathbf{vec}(U(\omega I - HA(x)H) + (\omega I - HA(x)H)U) \end{pmatrix}$$

and where $U = \mathbf{mat}(u)$, $H = \mathbf{mat}(h)$, and

$$\mathbf{mat} : \Re^{n(n+1)/2} \longrightarrow S\Re^{n\times n}$$

denotes the inverse of **vec**. Note that we now explicitly choose $H = H^T$.

Our algorithm essentially amounts to applying Newton's method to solve the nonlinear system of equations $F(z) = 0$.

Algorithm Let $\epsilon > 0$ be very small, e.g. ten times the machine precision. Let $z^0 = (x^0, \omega^0, h^0, u^0)$ with $U^0 = \mathbf{mat}(u^0) \succeq 0$, $H^0 = \mathbf{mat}(h^0) = B(x^0)^{-1/2}$, and $\omega^0 = \lambda_1(A(x^0), B(x^0)) + \epsilon$. For $k = 0, 1, \ldots$ do

1. Solve
$$F'(z^k)(\Delta z) = -F(z^k)$$
for $\Delta z = (\Delta x, \Delta\omega, \Delta h, \Delta u)$, where $F'(z)$ denotes the Jacobian matrix of F at z.
2. Update $u^{k+1} = u^k + \Delta u$, $x^{k+1} = x^k + \Delta x$. Compute the smallest eigenvalue θ_n of $U = \mathbf{mat}(u^{k+1})$. If $\theta_n < \epsilon$, replace u^{k+1} by $u^{k+1} + (\epsilon - \theta_n)\mathbf{vec}(I)$.
3. Let $h^{k+1} = \mathbf{vec}(B(x^{k+1})^{-1/2})$ and let $\omega^{k+1} = \lambda_1(A(x^{k+1}), B(x^{k+1})) + \epsilon$, and set $z^{k+1} = (x^{k+1}, \omega^{k+1}, h^{k+1}, u^{k+1})$.

Thus x and u are updated using the Newton step, without performing any line search. Then the resulting matrix $U = \mathbf{mat}(u)$ is shifted, if necessary, by a suitable multiple of the identity to ensure that the dual matrix U remains positive semidefinite. The choice of ω guarantees that the primal matrix remains positive semidefinite at every step. $H = \mathbf{mat}(h)$ is equal to the symmetric square root of $B(x)$ throughout the algorithm.

In primal-dual algorithms for LP or SDP one usually performs some form of a line search on the Newton step in order to force the variables to remain in the positive cone (LP) or positive semidefinite cone (SDP). The convergence rate of the resulting algorithm is very sensitive to the particular scheme that used in the line search [10,14]. We have found that the use of the full Newton step followed by a shift, if necessary, to be very efficient: quadratic convergence was achieved in every test problem.

5. Numerical Results

The algorithm was implemented in *Matlab*. We observe fast local convergence generically, at a quadratic rate. We show results in Table 1 for a random problem with $n = 10$ and $m = 15$, and with the pencil $(A(x), B(x))$ depending affinely on the parameters x, i.e.

$$A(x) = A_0 + \sum_{i=1}^{m} x_i A_i, \quad B(x) = B_0 + \sum_{i=1}^{m} x_i B_i$$

with $A_i, B_i \in S\Re^{n\times n}$ for $0 \leq i \leq m$. The matrix B_0 was set to the identity and $A_0, \ldots, A_{15}$ and $B_1, \ldots, B_{15}$ were generated using the *Matlab* `rand` function, and symmetrizing. The matrices $B_1, \ldots, B_{15}$ were then scaled by a factor 0.05 in order to ensure the existence of a reasonable-sized domain with $B(x) \succeq 0$. Finally, x^0 and U^0 were initialized with random data and U^0 was shifted so that $U^0 \succeq 0$. The total number of flops for this example was about 25 million, which is less than the number of flops for a single iteration of the method of centers, used by [2,3,6], as implemented by the authors. The data is available from the authors upon request.

Iteration	$\|\Delta z\|$	$\|F(z)\|$	ω
1	6.231e+01	1.038e+02	1.617e+00
2	1.852e+01	8.697e+01	1.580e+00
3	6.520e+00	1.712e+00	5.012e-01
4	1.564e+00	6.571e-01	2.057e-01
5	1.068e+00	1.388e-01	2.389e-01
6	4.372e-01	1.558e-01	1.600e-01
7	2.877e-01	4.013e-02	1.478e-01
8	1.163e-01	1.422e-01	1.428e-01
9	6.037e-02	3.321e-03	1.423e-01
10	2.214e-02	5.812e-04	1.422e-01
11	4.040e-03	8.105e-05	1.422e-01
12	1.437e-04	2.737e-06	1.422e-01
13	1.818e-07	3.470e-09	1.422e-01
14	2.897e-13	2.924e-14	1.422e-01

Table 1: Sample Output from the Algorithm

Acknowledgement. The work of the second author was supported in part by National Science Foundation Grant CCR-9101649.

References

[1] F. Alizadeh, "Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization," SIAM J. Optim., to appear.

[2] J.-P. A. Haeberly and M. L. Overton, "A hybrid algorithm for optimizing eigenvalues of symmetric definite pencils," SIAM J. Matrix Anal. Appl., 14 (1994), to appear.

[3] S. Boyd and L. El Ghaoui, "Method of centers for minimizing generalized eigenvalues," Linear Algebra Appl., to appear.

[4] R. W. Freund and F. Jarre, "An interior-point method for convex fractional programming," AT&T Bell Labs Numerical Analysis Manuscript 93–03, Murray Hill, NJ, 1993. Submitted to Math. Programming.

[5] Y. Nesterov and A. Nemirovsky, "Interior-Point Polynomial Algorithms in Convex Programming," SIAM, Philadelphia, 1994.

[6] B. Nekooie and M. K. H. Fan, A second-order interior point method for solving linear matrix inequality problems," manuscript, Dept. of Elec. Eng., Georgia Inst. Tech, 1993. Submitted to SIAM J. Control Optim.

[7] I. J. Lustig, R. E. Marsten and D. F. Shanno, "Interior Point Methods for Linear Programming: Computational State of the Art," ORSA Journal of Computing, 6:1–14, 1994.

[8] L. Vandenberghe and S. Boyd, "A primal-dual potential reduction method for problems involving matrix inequalities," Math. Programming Series B, to appear.

[9] F. Rendl, R. Vanderbei and H. Wolkowicz, "Max-Min Eigenvalue Problems, Primal-Dual Interior Point Algorithms, and Trust Region Subproblems," Technical Report CORR 93-30, University of Waterloo, 1993. To appear in Proceedings of NATO Conference on Nonlinear Programming, Il Ciocco, Italy, Sept. 1993.

[10] F. Alizadeh, J.-P. Haeberly, and M. L. Overton, "A New Primal-Dual Interior-Point Method for Semidefinite Programming," Computer Science Dept. Report 659, New York University. To appear in Proceedings of Fifth SIAM Conf. on Applied Linear Algebra, Snowbird, Utah, June 1994.

[11] M. L. Overton, "On minimizing the maximum eigenvalue of a symmetric matrix," SIAM J. Matrix Anal. Appl., 9:256–268,1988.

[12] M. L. Overton, "Large-scale optimization of eigenvalues," SIAM J. Optim. 2:88–120, 1992.

[13] M. L. Overton and R. S. Womersley, "Second derivatives for optimizing eigenvalues of symmetric matrices," Computer Science Dept. Technical Report 627, New York University, 1993. SIAM J. Matrix Anal. Appl, to appear.

[14] Y. Zhang, R.A. Tapia, and J.E. Dennis, "On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms," SIAM Journal on Optimization, 2:304–324, 1992.