A Simple Application of the Homotopy Method to Symmetric Eigenvalue Problems

Moody T. Chu
Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27650

Submitted by George P. Barker

ABSTRACT

The homotopy method is used to find all eigenpairs of symmetric matrices. A special homotopy is constructed for Jacobi matrices. It is shown that there are exactly n distinct smooth curves connecting trivial solutions to desired eigenpairs. These curves are solutions of a certain ordinary differential equation with different initial values. Hence, they can be followed numerically. Incorporated with sparse matrix techniques, this method might be used to solve eigenvalue problems for large scale matrices.

1. INTRODUCTION

Solving a symmetric eigenvalue problem

$$Ax = \lambda x \tag{1.1}$$

can be thought of as solving a nonlinear algebraic equation

$$f(x, \lambda) = 0, \tag{1.2}$$

where $f: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}$ is defined by

$$f(x, \lambda) = \left(Ax - \lambda x, \frac{1 - x^T x}{2} \right). \tag{1.3}$$

From this point of view, many well-developed methods can then be employed to find zeros of this f. For example, if the spectrum $\sigma(A)$ of A is simple, as will be assumed henceforth, then the classical Newton's method and its many

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

improved modifications are particularly well suited for solving (1.2), since its higher Frechét derivatives can easily be determined—the second derivative is constant and higher derivatives vanish. For a general discussion of this approach, see, for example, [6] and [11]. Unfortunately, Newton's method can converge (if it ever converges) to only one zero at a time. In order to obtain all n eigenpairs of A, we have to restart the iteration by making n suitable guesses. In this paper we are interested in applying the homotopy method to solve the problem (1.2). As will be shown in the sequel, one of the advantages is the existence of n disjoint smooth curves in $\mathbb{R}^n \times \mathbb{R}$ each of which leads from an obvious starting point to the desired eigenpair (x, λ) of (1.1). Furthermore, together with the large scale matrix techniques, this method can be used to solve eigenvalue problems for sparse matrices.

Homotopy methods for finding zeros are known variously as "continuation methods" or "Davidenko's methods." The basic idea is to construct a homotopy from a trivial map to the one of interest. Under suitable conditions, a curve starting from the trivial solution will then lead us by a smooth path to the desired solution. This method in spirit is the same as the degree theory. But in practice it becomes very attractive because of its global and probabilistic features. For a general discussion of the underlying theory and some of its numerical treatments, see, for example, [1, 2, 5, 7, 8, 9, 12].

The difficulty encountered in applying the homotopy method is the selection of an appropriate homotopy equation so that (1) the existence of a curve connecting the trivial solution and the desired solution is assured and (2) the numerical work in following this curve is at reasonable cost. Toward this end, one possible approach is to view (1.3) as a system of $n+1$ quadratic polynomials in $n+1$ unknowns. Then the special homotopy suggested in [3] seems applicable for solving (1.2). One realizes immediately, however, that there is a great waste in doing so, since there are at least $2^{n+1}-n$ curves diverging to infinity. In this paper a much simpler homotopy is constructed. We show that there are exactly n disjoint curves leading us to the n eigenpairs. These curves are characterized by an explicit ordinary differential equation with distinct initial values. Thus they can be followed easily by any available ODE software solvers.

We begin in the next section with a description of this particular homotopy equation. Then we show why this homotopy will work. Some comments about the applicability of this method are made in Section 3.

2. METHOD AND THEOREMS

Through a standard tridiagonalization process, we may assume, without loss, that the matrix A is a Jacobi matrix with nonzero off-diagonal elements.

Let D be an arbitrary diagonal matrix with distinct elements. Construct the homotopy equation $H:\mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}$ as follows:

$$H(x,\lambda,t) = \left([D + t(A-D)]x - \lambda x, \frac{1 - x^T x}{2} \right). \tag{2.1}$$

We claim

LEMMA 1. *The point* $0 \in \mathbb{R}^n \times \mathbb{R}$ *is a regular value for* H. *In other words, for each* $(x,\lambda,t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ *such that* $H(x,\lambda,t) = 0$ *the Jacobian matrix* $D_{(x,\lambda,t)}H$ *always has rank* $n+1$.

Proof. Observe that

$$D_{(x,\lambda,t)}H = \begin{bmatrix} D + t(A-D) - \lambda & -x & (A-D)x \\ -x^T & 0 & 0 \end{bmatrix}. \tag{2.2}$$

If $H(x,\lambda,t) = 0$, then

$$[D + t(A-D)]x = \lambda x \tag{2.3}$$

and

$$x^T x = 1.$$

Equation (2.3) implies that the vector x is orthogonal to the rows of $D + t(A - D) - \lambda$. Since $D + t(A-D) - \lambda$ also has simple spectrum [4, Lemma 6.1], the first n columns of $D_{(x,\lambda,t)}H$ are linearly independent. By symmetry, we also know the $(n+1)\times(n+1)$ square matrix

$$D_{(x,\lambda)}H = \begin{bmatrix} D + t(A-D) - \lambda & -x \\ -x^T & 0 \end{bmatrix} \tag{2.5}$$

has rank $n+1$. This finishes the proof. ∎

Indeed, one even knows that

$$\sigma(D_{(x,\lambda)}H) = \sigma(D + t(A-D) - \lambda) \cup \{1, -1\} - \{0\}. \tag{2.6}$$

The following lemma is a standard result from the differential topology [10].

LEMMA 2. *The set* $\Gamma = \{(x, \lambda, t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : H(x, \lambda, t) = 0\}$ *is a one dimensional smooth manifold. Hence, each component is a smooth curve and is either diffeomorphic to a circle or an interval.*

Along each component, we may take the derivative with respect to the arc length s. The set Γ is then characterized by

$$D_{(x,\lambda)}H \cdot \begin{bmatrix} \frac{dx}{ds} \\ \frac{d\lambda}{ds} \end{bmatrix} + D_t H \cdot \frac{dt}{ds} = 0. \tag{2.7}$$

Notice that Γ will never turn back, i.e. $dt/ds \neq 0$, since otherwise $D_{(x,\lambda)}H$ would be singular. Thus Γ can be parametrized by the variable t and (2.7) becomes

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{d\lambda}{dt} \end{bmatrix} = \begin{bmatrix} D + t(A - D) - \lambda & -x \\ -x^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} -(A - D)x \\ 0 \end{bmatrix} \tag{2.8}$$

Let $\Gamma_0 = \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : (H(x, \lambda, t) = 0 \text{ for some } t \in [0, 1]\}$. It is important to know whether the curve Γ will diverge to infinity or not. For this we claim

LEMMA 3. *The set* Γ_0 *is bounded. Together with the fact that* $dt/ds \neq 0$, *the curve, therefore, always connects the trivial solution of* $H(x, \lambda, 0) = 0$ *to the desired solution of* $H(x, \lambda, 1) = 0$.

Proof. For any $(x, \lambda, t) \in \Gamma$, it is clear that

$$\|(x(t), \lambda(t))\| \leqslant 1 + |\lambda(t)|. \tag{2.9}$$

By (2.3), $|\lambda(t)|$ is bounded by

$$|\lambda(t)| \leqslant \|D + t(A - D)\| \leqslant \|D\| + \|A\|$$

if $t \in [0, 1]$. In fact we can get bounds trivially by Gerschgorin's theorem.

3. APPLICABILITY

We have shown that Γ consists of exactly n curves which are solutions of the differential equation (2.8). Hence, these curves may be followed numerically by using any of the available differential equation solvers. To be more specific, to follow the ith component of Γ, we start from the trivial solution

$$\begin{bmatrix} x(0) \\ \lambda(0) \end{bmatrix} = \begin{bmatrix} e_i \\ d_i \end{bmatrix}, \tag{3.1}$$

where e_i is the standard ith unit vector and d_i is the ith element of D. The step-by-step evaluation of the tangent vector can be obtained by solving the system

$$\begin{bmatrix} D + t(A-D) - \lambda & -x \\ -x^T & 0 \end{bmatrix} \begin{bmatrix} \dfrac{dx}{dt} \\ \dfrac{d\lambda}{dt} \end{bmatrix} = \begin{bmatrix} -(A-D)x \\ 0 \end{bmatrix} \tag{3.2}$$

using any typical linear equation solvers. It is worth noting that the standard LU decomposition, for example, will produce one fill-in only at the (n,n) position if A is a Jacobi matrix, so the method is particularly attractive for solving large scale tridiagonal eigenvalue problems.

It should be noted also that the Jacobi structure of the matrix A is only a sufficient condition for Lemma 1. This condition may be rephrased as "chosing D so that the matrix $(1-t)D + tA$ always has simple spectrum for each $t \in [0,1]$." Alternatively, all we require is the full rank of the matrix $D_{(x,\lambda,t)}H$, and then we can use (2.7) instead (2.8). Evidently the sparse matrix techniques can be incorporated in any of these cases.

REFERENCES

1. E. Allgower and K. Georg, Simplicial and continuation methods for approximating fixed points and solutions to systems of equations, *SIAM Rev.* 22:28–85 (1980).
2. S. N. Chow, J. Mallet-Paret, and J. A. Yorke, Finding zeros of maps: Homotopy methods that are constructive with probability one, *Math. Comp.* 32:887–899 (1978).
3. ———, A homotopy method for locating all zeros of a system of polynomials, in functional differential equations and approximation of fixed points, *Lecture Notes Math.* 730:77–88 (1979).

4. P. Deift, F. Lund, and E. Trubowitz, Nonlinear wave equations and constrained harmonic motion, *Comm. Math. Phys.* 74:141–188 (1980).
5. M. T. Chu, On a numerical treatment for the curve tracing of the homotopy method, Numer. Math., to appear.
6. L. Collatz, *Functional Analysis and Numerical Mathematics,* Academic, New York, 1966.
7. C. B. Garcia and F. J. Gould, A theorem on homotopy paths, *Math. Oper. Res.* 3:282–289 (1978).
8. K. Georg, Numerical intepration of the Davidenko equation, in *Numerical Solution of Nonlinear Equations*, Lecture Notes in Math., Vol. 878, 1980, pp. 128–161.
9. M. Hirsch and S. Smale, On algorithms for solving $f(x)=0$, *Comm. Pure Appl. Math.* 32:281–312 (1979).
10. J. Milnor, *Topology from the Differentiable Viewpoint*, Univ. of Virginia Press, 1965.
11. H. Unger, Nichtlineare Behandling von Eigenwertaufgaben, Z. *Angew. Math. Mech.* 30:281–282 (1950).
12. H. J. Wacker, *Continuation Methods*, Academic, New York, 1978.

Received 20 March 1983