Vol. 1, No. 2
Winter, 1990

AN ALGORITHM FOR FINDING THE EIGEN-PAIRS OF A SYMMETRIC MATRIX

M. Mohseni Moghadam

Department of Mathematics, Faculty of Science, University of Kerman, Kerman, Iran

Keywords: Eigenvalue, Homotopy, Eigenvector

Abstract

The purpose of this paper is to show that ideas and techniques of the homotopy continuation method can be used to find the complete set of eigenpairs of a symmetric matrix. The homotopy defined by Chow, Mallet-Paret and York [1] may be used to solve this problem with 2^{n+1}-n curves diverging to infinity which for large n causes a great inefficiency. M. Chu [2] introduced a homotopy equation to solve this problem. In this method it is necessary to follow 2n curves to handle the problem. Our method is based on a special homotopy system of equations which consists of exactly n distinct smooth curves and connects trivial solution to desired eigenpairs. It is important that in our method we avoid finding explicitly the coefficient of the characteristic equation, as all experienced practitioners are aware of the large error that may result from the use of the approximate coefficients of the characteristic polynomial.

(1) Introduction

The eigenpair (eigenvector, eigenvalue) problem for a square matrix $A \in \mathbb{R}^{n^2}$ is that of determining a scalar λ and a vector x such that

$$Ax=\lambda x, x\neq 0 \qquad (1.1)$$

The problem is clearly nonlinear since both λ and x are unknown. Since the eigenvalues are the n roots of the characteristic equation

$$\det(A-\lambda I)=P(\lambda)=0 . \qquad (1.2)$$

They can be found without reference to any of the eigenvectors. For a given eigenvalue λ, the corresponding eigenvector is a nontrivial solution of the linear system $Ax=\lambda x$.

This paper is concerned with the homotopy continuation method for calculating the complete set of eigenpairs of a symmetric matrix, and we avoid finding explicitly the coefficients of $P(\lambda)$ in order to determine the eigenvalues. Instead, a special homotopy is introduced and we shall prove that there are exactly n distinct smooth curves which connect trivial solutions to the desired eigenpairs. In fact, these curves are solutions of certain ordinary differential equations with different initial values, and hence they can be followed numerically by any ordinary differential equations solver.

We emphasize the practical importance of not finding explicitly the coefficients of $P(\lambda)$ in order to evaluate the polynomial. All experienced practitioners are aware of the large error that may result from the use of the approximate coefficients of $P(\lambda)$ for calculation of the zeros of the characteristic polynomials.

(2) The Algorithm

We restrict our discussion to the symmetric eigenpair problem

$$Ax=\lambda x, X\neq 0 . \qquad (2.1)$$

Although through a standard tridiagonalization we may assume, without loss of generality, that the matrix A is a

Jacobian matrix with nonzero off-diagonal elements.

The eigenpair problem can be thought of as solving a nonlinear algebraic equation

$$f(x,\lambda)=0\,, \quad (2.2)$$

where

$$f:\mathbb{R}^n\times\mathbb{R}\longrightarrow\mathbb{R}^n$$

is defined by

$$f(x,\lambda)=Ax-\lambda x.$$

There are many well-developed methods which can be used to find the non-trivial solutions of $f(x,\lambda)=0$. For example, suppose that the spectrum $\sigma(A)$ is simple. Then the classical Newton's method and its many improved modifications are particularly well suited for solving (2.2). Since its higher Frechet derivatives can easily be determined, the second derivative is constant and higher derivatives vanish. For a detailed discussion of this approach, see [3]. Unfortunately there are some disadvantages in using the Newton method. One of them being that Newton's method can converge (if it ever converges) to only one eigenpair at a time. That is in order to compute all n eigenpairs of A, we have to restart the iteration by making n suitable guesses. One possible approach to solving this problem by the homotopy continuation method is to view (2.2) as a system of n+1 quadratic polynomials in n+1 unknowns. Then the special homotopy defined by Chow, Mallet-Paret and York [1] is applicable for solving (2.2). However, there are at least 2^{n+1}-n curves diverging to infinity which causes a great inefficiency (particulary for large n). Another approach presented by M.Chu to solve this problem is [2]. He defined a homotopy:

$$H:\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\longrightarrow\mathbb{R}^n\times\mathbb{R}$$

by

$$H(x,\lambda,t)=([D+t(A-D)-\lambda I]x,\tfrac{1}{2}(x^tx-1)), \quad (2.3)$$

where D is an arbitrary diagonal matrix with distinct elements. Applying this homotopy has the following disadvantages; a) if we follow the n distinct curves suggested by M.Chu we may not get all eigenpairs, since two of these curves may link into a pair of eigenpairs of the form (x,λ) and $(-x,\lambda)$, (so they actually represent one eigenpair), b) to get all eigenpairs we actually must follow 2n distinct curves rather than n curves.

In order to remedy this problem and solve (2.2) at a reasonable cost, a special homotopy is constructed as follows:

Let D be an arbitrary diagonal matrix with distinct elements on its diagonal. Construct the homotopy equation

$$H:\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\longrightarrow\mathbb{R}^n\times\mathbb{R}$$

defined by

$$H(x,\lambda,t)$$

$$=([(1-t)D+tA-\lambda I]x\,,\varepsilon\sum_{i=1}^{n}x_i\;-t(x^tx)^{\frac{1}{2}}-1), \quad (2.4)$$

where ε is a small positive number. It is clear that vectors

$$\begin{bmatrix} x(0)\\ \lambda(0)\end{bmatrix}=\begin{bmatrix} e_i/\varepsilon\\ d_i\end{bmatrix} i=1,2,\ldots,n$$

are the eigenpairs of $H(x,\lambda,0)$, where e_i is the standard ith unit vector and d_i is the i'th element of the diagonal matrix D. We should mention here that the crucial step in applying the homotopy continuation method is the construction of an appropriate homotopy, such as (2.4) so that i) the existence of a curve connecting the trivial solution and desired solution is assured and ii) the numerical work in following this curve has a reasonable cost.

In the next section we shall show that the homotopy equation (2.4) guarantees the existence of n distinct smooth curves, each of them leading from an obvious starting point to a desired eigenpair. Furthermore, if a certain curve links to an eigenpair (x,λ), then there is no other curve that may link to $(-x,\lambda)$. These curves are characterized by an explicit ordinary differential equation with distinct initial values, and hence they can be easily followed by any ordinary differential equations solver. Coupled with the large scale matrix techniques, this method can be used to solve eigenvalue problems for sparse matrices [5].

(3) Theorems

In this section we present some theorems which serve as a theoretical basis for our algorithm.

Theorem (3.1)

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

Proof

Let $(\bar{x},\bar{\lambda},\bar{t})\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}$ and $H(\bar{x},\bar{\lambda},\bar{t})=0$. Observe that

$$D_{(\bar{x},\bar{\lambda},\bar{t})}H=$$

$$\begin{bmatrix} (1-\bar{t})D+\bar{t}A-\bar{\lambda}I & -\bar{x} & (A-D)\bar{x} \\ (\epsilon-t(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_1,\ldots,\epsilon^{-t}(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_n) & 0 & -(\bar{x}^t\bar{x})^{-\frac{1}{2}} \end{bmatrix} \quad (3.1)$$

Since $H(\bar{x},\bar{\lambda},\bar{t})=0$, we have

$$\epsilon\sum_{i=1}^{n} \bar{x}_i-\bar{t}(\bar{x}^t\bar{x})^{\frac{1}{2}}=1 \quad (3.2)$$

and

$$((1-\bar{t})D+\bar{t}A-\bar{\lambda}I).\bar{x}=0. \quad (3.3)$$

We claim that the $(n+1)x(n+1)$ matrix

$$D_{(\bar{x},\bar{\lambda})}H=\begin{bmatrix} (1-\bar{t})D+\bar{t}A-\bar{\lambda}I & -\bar{x} \\ (\epsilon-\bar{t}(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_1,\ldots,\epsilon-\bar{t}(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_n) & 0 \end{bmatrix}$$

is of full rank. Since otherwise there exists a vector $(y,\mu)^t\neq 0$ with $y\epsilon\mathbb{R}^n$, and $\mu\epsilon\mathbb{R}^n$ such that

$$D(\bar{x},\bar{\lambda})H.\begin{pmatrix} y \\ \mu \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3.4)$$

thus

$$((1-\bar{t})D+\bar{t}A-\bar{\lambda}I).y-\mu\bar{x}=0 \quad (3.5)$$

This implies

$$\bar{x}^t.((1-\bar{t})D+\bar{t}A-\bar{\lambda}I).y=\mu\bar{x}^t\bar{x}.$$

Since A is symmetric, and $\bar{x}$ is orthogonal to the rowspace of the $(1-\bar{t})D+\bar{t}A-\bar{\lambda}I$, we have

$$\mu\|\bar{x}\|_2^2=\mu\bar{x}^t\bar{x}=(\bar{x}^{-t}.\ (1-\bar{t})D+\bar{t}A-\bar{\lambda}I).y)^t$$
$$=y^t.((1-\bar{t})D+\bar{t}A-\bar{\lambda}I).\bar{x}=0.$$

This implies $\mu=O$. Therefore

$$((1-\bar{t})D+\bar{t}A-\bar{\lambda}I).y=0. \quad (3.6)$$

Since the matrix $(1-\bar{t})D+\bar{t}A-\bar{\lambda}\ I$ has a simple spectrum ([4], Lemma 6.1), we conclude that the matrix $B=(1-\bar{t})D+\bar{t}A-\bar{\lambda}I$ has a set of orthogonal eigenvectors say, z_1, $z_2,\ldots,z_n$, with corresponding eigenvalues $\delta_1,\delta_2,\ldots,\delta_n$. Let

$$\bar{x}=\alpha_1z_1+\alpha_2z_2+\ldots+\alpha_nz_n,$$
$$y=\beta_1z_1+\beta_2z_2+\ldots+\beta_nz_n.$$

Then by (3.3), (3.6) and orthogonality we get

$$\alpha_i=\beta_i\delta \qquad i=1,2,\ldots,n.$$

Therefore $y=\delta\bar{x}$ for some δ. Substituting $y=\delta\bar{x}$ and $\mu=0$ in (3.4)

we get

$$\begin{bmatrix} (1-\bar{t})D+\bar{t}A-\bar{\lambda}I & -\bar{x} \\ \epsilon-\bar{t}(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_1,\ldots,\epsilon-\bar{t}(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_n & 0 \end{bmatrix}\begin{bmatrix} \delta\bar{x} \\ 0 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Thus by (3.2) we have,

$$0=\delta(\epsilon\sum_{i=1}^{n}\bar{x}_i-\bar{t}(\bar{x}^t\bar{x})^{\frac{1}{2}})=\delta. \quad (3.7)$$

Hence $D_{(\bar{x},\bar{\lambda})}H$ is of rank $n+1$. This completes the proof.

Remark. We have restricted our discussion to the Jacobian structure of matrix A. This is needed only as a sufficient condition for Theorem (3.1). This condition may be rephrased as «choosing D so that the matrix (1-t) D+tA has a simple spectrum for any $t\epsilon[0,1]$». Overall, it is only needed that the matrix $D_{(x,\lambda,t)}H$ be of full rank for any $(x,\lambda,t)\ \epsilon\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}$ with $H\ (x,\lambda,t)=0$. Apparently the sparse matrix techniques can be incorporated in any of these cases.

As is the usual procedure of the homotopy continuation method we start from a trivial solution of $H(.,.,0)$ at $t=0$, and follow the generated path as t increases from zero to one. We hope the trivial eigenpairs deform into the eigenpairs of the original matrix A. Hence, we would be able to follow the n distinct connected paths from the trivial system to the original problem. In order to assure that this process works, we prove the following:

Theorem (3.2) Let us define

$$\Gamma=\{(x,\lambda,t)\ \epsilon\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}:H\ (x,\lambda,t)=0\}.$$

Then

a) Γ is a one dimensional smooth manifold

b) as t increases, the curve Γ will never turn back.

Proof: Part (a) is in fact a standard result from the differential topology [6]. That is, a repeat use of the Implicit Function Theorem implies that Γ consists of one dimensional manifold.

In order to prove (b), let Γ be parameterized with a parameter θ. Along each component, we may take the derivative with respect to the parameter θ. The set Γ is then characterized by

$$D_{(t)}H.\frac{dt}{d\theta}+D_{(x,\lambda)}H.\begin{bmatrix}\frac{dx}{d\theta}\\ \frac{d\lambda}{d\theta}\end{bmatrix}=0. \quad (3.8)$$

We claim $\frac{dt}{d\theta}\neq 0$, since otherwise

$$D_{(x,\lambda)}H.\begin{bmatrix}\frac{dx}{d\theta}\\ \frac{d}{d\theta}\end{bmatrix}=0.$$

Hence $D_{(x,\lambda)}H$ would be singular. This completes the proof. Now consider starting a path at $t = 0$. Since $t^{\cdot}\neq 0$. Without loss of generality, we may assume $\dot{t}>0$. So as we increase the parameter θ, $t(\theta)$ cannot reverse. The following lemma shows that as t tends to one the curve Γ remains bounded. In other words for any $0<t_O<1$, the set

$$\Gamma_O=\{(x,\lambda)\in \mathbb{R}^n\times\mathbb{R}: H(x,\lambda,t)=0 \text{ for some } t\in[t_O,1]\} \quad (3.9)$$

is bounded.

Lemma (3.3) The set Γ_O is a bounded set.

Proof: By equation (3.3)

$$|\lambda(t)|\|x\|=\|[1\text{-}t)D+tA]x\|\leq\|(1\text{-}t)D+tA\|\;\|x\|$$
$$\leq(\|D\|+\|A\|)\|x\|$$

Hence

$$|\lambda(t)|\leq\|D\|+\|A\| \quad (3.10)$$

Also for any $(x,\lambda)\epsilon\Gamma_O$, equation (3.2) implies

$$t_o\|x\|\leq\|x\|=|1+\varepsilon\sum_{i=1}^{n}x_i|$$
$$\leq 1+\varepsilon\sum_{i=1}^{n}|x_i|=1+\varepsilon\|x\|_1$$
$$\leq 1+\varepsilon\sqrt{n}\|x\|.$$

Thus for ε small enough such that $\varepsilon\sqrt{n}<t_O$, we have

$$\|x\|\leq\frac{1}{t_O\text{-}\varepsilon\sqrt{n}}. \quad (3.11)$$

Therefore Γ_O is bounded. (Here by $\|.\|$ we mean $\|.\|_2$.)

Part (b) of Theorem (3.2) implies that Γ will never turn back. Thus Γ can be parametrized by the variable t. Then (3.8) becomes

$$\begin{bmatrix}\frac{dx}{dt}\\ \frac{d\lambda}{dt}\end{bmatrix}=\begin{bmatrix}(1-t)D+tA-\lambda I & -x\\ \epsilon-t(x^tx)^{-\frac{1}{2}}x_1,\ldots\epsilon-t(x^tx)^{-\frac{1}{2}}x_n & 0\end{bmatrix}^{-1}\begin{bmatrix}(D-A)x\\ -(x^tx)^{\frac{1}{2}}\end{bmatrix}$$

$$\begin{bmatrix}x(0)\\ \lambda(0)\end{bmatrix}=\begin{bmatrix}e_i/\epsilon\\ di\end{bmatrix} \quad i=1,2,\ldots,n. \quad (3.12)$$

These differential equations may be solved by any ordinary differential equations solver. Over all for each $1\leq i\leq n$ the numerical solution of (3.12) at t = 1 is an approximation for an eigenpair of the matrix A. The following lemma assures that among the computed eigenpairs we will not have a pair of eigenpairs of the form $[x,\lambda]^t$ and $[-x,\lambda]^t$.

Lemma (3.4) There is no pair of eigenpairs of the form

$$[x,\lambda]^t,\quad [-x,\lambda]^t.$$

Proof: Since otherwise both vectors satisfy equation (3.2) with t=1. Subtracting the two resulting equations gives $\sum_{i=1}^{n}x_i=0$. Substituting this in (3.2) yields $\|x\|_2=-1$, a contradiction.

(4) The Algorithm

We have seen that the set

$$\Gamma=\{(x,\lambda,t):H(x,\lambda,t)=0\}. \quad (4.1)$$

consists of exactly n curves, and each of them is the solution of the differential equations (3.12). We proceed as follows:

I) Choose any diagonal matrix D with distinct elements on its diagonal.

II) Set i = 0.

III) Set i=i+1, and take

$$\begin{bmatrix}x(o)\\ \lambda(o)\end{bmatrix}=\begin{bmatrix}ei/\epsilon\\ di\end{bmatrix} \quad (4.2)$$

IV) Solve the differential equations (3.12) with the initial value (4.2) by using any ordinary differential equations solver.

V) At t=1, write down the computed values for the vector

$$\begin{bmatrix}x(1)\\ \lambda(1)\end{bmatrix}.$$

(This is the **ith** eigenpair of matrix A).

VI)If i <n, go to III,

VII) If i=n stop.

References

1. Chow., S.N., Mallet-Paret, J.and Yorme, J.A.,A"Homotopy Method for Locating all Zeros of a System Polynomials, Functional Differential Equations and Approximation of Fixed Points, Springer Lecture Notes in Mathematics, **No.730,** pp.77-88. (1978).
2. Chu,M.T., A Simple Application of the Homotopy Method Symmetric Eigenvalue Problem, (to appear).
3. Collatz, L.Functional Analysis and Numerical Mathematics, Acd. Press. New York. (1966).
4. Deift, P.,Lund, F.and Trubowitz, E.,Nonlinear Wave Equation and Constrained Harmonic Motion, *Comm. Math. Phys.* **74,** pp.141-188. (1980).
5. Garcia, C.B. and Zangwill, W.I., An Approach to Homotopy and Degree Theory, Math. of Operations Research, **Vol.4,** pp. 390-405.(1979).
6. Milnor, J., Topology from the Differential View Point, *Univ. Virginia Press,* VA. (1965).