Numerical solution of the inverse algebraic eigenvalue problem

By Z. Bohte*

The inverse algebraic eigenvalue problem is formulated and a numerical method for its solution explained. The convergence of the method is considered and a numerical example given.

1. Introduction

The practical inverse algebraic eigenvalue problem arose in physical chemistry in the study of molecular structures. For the numerical solution of this problem a number of iterative techniques has been used without sufficient theoretical consideration. We refer to the most recent paper on the subject of Toman and Pliva (1966) where other references can be found. We believe the problem has not attracted the attention of mathematicians yet.

As the problem is in general non-linear it is difficult to compare different methods. It may happen that they all fail to yield a result but in our experience the method which we are going to explain and which has been most popular in practice has at least theoretical advantages.

2. Formulation of the general problem

The general inverse algebraic eigenvalue problem can be formulated as follows.

Let

$$A = A(p_1, \ldots, p_n) \tag{1}$$

be a given n by n matrix which is an analytic function of n parameters $p_1, \ldots, p_n$. The problem is to find the values of parameters $p_1^*, \ldots, p_n^*$ such that the matrix $A(p_1^*, \ldots, p_n^*)$ possesses given eigenvalues $\lambda_1^*, \ldots, \lambda_n^*$.

In general the eigenvalues of (1) are rather complicated non-linear functions of parameters. Essentially the problem requires a solution of the non-linear system of equations

$$\lambda_i(p_1, \ldots, p_n) = \lambda_i^* \quad (i = 1, \ldots, n) \tag{2}$$

where the functions λ_i are not explicitly known. We know, of course, how to find the values of λ_i if the parameters are given. Usually, a good initial approximation to parameters is available, so it is possible to apply various iterative numerical methods for the solution of the problem. Some of these methods require additional information about the functions λ_i.

In all practical cases we have encountered, the matrix A has been a symmetric matrix and a linear function of parameters:

$$A = A_0 + \sum_{k=1}^{n} p_k A_k, \quad A_k^T = A_k \quad (k = 0, \ldots, n) \tag{3}$$

and we shall limit our considerations to this case.

* *University of Surrey, London, S.W.2.*

3. Properties of the eigensystem

F. Rellich (1936) proved that the eigenvalues and eigenvectors of a symmetric matrix $A(z)$ can be expressed as convergent power series of z provided $A(z)$ is an analytic function of z. He also stated a similar theorem in the case of more parameters. For the aim of this paper we need a more explicit result in our particular case. We shall prove the following:

Theorem. Let A be a real symmetric n by n matrix and let it be a linear function of m real parameters

$$A = A_0 + \sum_{i=1}^{m} p_i A_i.$$

If λ_k is a simple eigenvalue of the matrix A_0 and x_k the associated eigenvector then the corresponding eigenvalue μ_k and the associated eigenvector y_k of A can be expressed as convergent power series of $p_1, \ldots, p_m$ in some neighbourhood of $p_1 = \ldots = p_m = 0$.

PROOF. Let $\lambda_1, \ldots, \lambda_n$ and $x_1, \ldots, x_n$ form the eigensystem of the matrix A_0 where $\lambda_j \neq \lambda_k$, $j \neq k$ and $x_i^T x_j = \delta_{ij}$. Assume

$$\left.\begin{aligned} \mu_k &= \lambda_k + \sum_{r=1}^{\infty} \sum_{\Sigma i = r} \lambda_{i_1 \ldots i_m} p_1^{i_1} \ldots p_m^{i_m} \\ y_k &= x_k + \sum_{r=1}^{\infty} \sum_{\Sigma i = r} x_{i_1 \ldots i_m} p_1^{i_1} \ldots p_m^{i_m} \end{aligned}\right\} \tag{4}$$

where $\Sigma i = i_1 + \ldots + i_m$.

We shall prove that the series (4) converge in the "cube" $|p_i| \leqslant R$ and find a bound for R.

Substituting (4) into equation $Ay_k = \mu_k y_k$ we obtain

$$\begin{aligned} &(A_0 + \sum_{i=1}^{m} p_i A_i)(x_k + \sum_{r=1}^{\infty} \sum_{\Sigma i = r} x_{i_1 \ldots i_m} p_1^{i_1} \ldots p_m^{i_m}) \\ &= (\lambda_k + \sum_{r=1}^{\infty} \sum_{\Sigma i = r} \lambda_{i_1 \ldots i_m} p_1^{i_1} \ldots p_m^{i_m}) \\ &\qquad \times (x_k + \sum_{r=1}^{\infty} \sum_{\Sigma i = r} x_{i_1 \ldots i_m} p_1^{i_1} \ldots p_m^{i_m}). \end{aligned} \tag{5}$$

The constant terms in (5) are identically equal. Equating terms of $p_1^{i_1} \ldots p_m^{i_m}$ on both sides of (5) where $\Sigma i = r$ we obtain

$$\begin{aligned} &A_0 x_{i_1 \ldots i_m} + \sum_{j=1}^{m} A_j x_{i_1 \ldots i_m}^{(j)} = \sum_{t=0}^{r} \sum_{\Sigma k = t} \lambda_{k_1 \ldots k_m} x_{i_1 - k_1 \ldots i_m - k_m} \\ &= \lambda_k x_{i_1 \ldots i_m} + \lambda_{i_1 \ldots i_m} x_k + \sum_{t=1}^{r-1} \sum_{\Sigma k = t} \lambda_{k_1 \ldots k_m} x_{i_1 - k_1 \ldots i_m - k_m} \end{aligned} \tag{6}$$

where $x^{(j)}_{i_1 \dots i_m} = x_{i_1 \dots i_{j-1} \dots i_m}$ and $\lambda_{0 \dots 0} = \lambda_k$. Any coefficient x with any negative index is to be replaced by 0.

Equations (6) for $r = 1, 2, \dots$ can be solved successively. Assuming that all coefficients with the index sum less than r have been calculated we can determine $\lambda_{i_1 \dots i_m}$ and $x_{i_1 \dots i_m}$ in the following way.

Rearranging (6) we obtain

$$(A_0 - \lambda_k I)x_{i_1 \dots i_m} = \lambda_{i_1 \dots i_m} x_k + \sum_{t=1}^{r-1} \sum_{\Sigma k = t} \lambda_{k_1 \dots k_m} \times x_{i_1 - k_1 \dots i_m - k_m} - \sum_{j=1}^{m} A_j x^{(j)}_{i_1 \dots i_m}. \quad (7)$$

Expressing the unknown vector $x_{i_1 \dots i_m}$ in the form

$$x_{i_1 \dots i_m} = \sum_{s=1}^{n} b^{(s)}_{i_1 \dots i_m} x_s \quad (8)$$

and substituting into (7) we find

$$b^{(s)}_{i_1 \dots i_m} = x_s^T \Big(\sum_{t=1}^{r-1} \sum_{\Sigma k = t} \lambda_{k_1 \dots k_m} x_{i_1 - k_1 \dots i_m - k_m} - \sum_{j=1}^{m} A_j x^{(j)}_{i_1 \dots i_m} \Big) / (\lambda_s - \lambda_k) \quad (9)$$

for all $s \neq k$. Since the coefficient of x_k vanishes on the left-hand side of (7) the corresponding coefficient on the right-hand side must also vanish and this enables us to determine $\lambda_{i_1 \dots i_m}$. For reasons of simplicity we select at each step $b^{(k)}_{i_1 \dots i_m} = 0$ which means that $x_k^T x_{i_1 \dots i_m} = 0$ and so we obtain a simple formula for $\lambda_{i_1 \dots i_m}$.

$$\lambda_{i_1 \dots i_m} = \sum_{j=1}^{m} x_k^T A_j x^{(j)}_{i_1 \dots i_m}. \quad (10)$$

In order to prove the convergence of the series (4) we shall find a majorizing power series in one variable. Denoting

$$M = \max_{1 \leqslant i \leqslant m} ||A_i||, \quad d = (n-1)/\min_{s \neq k} |\lambda_s - \lambda_k|$$

we can easily obtain from (8), (9), and (10) the following bounds

$$|\lambda_{i_1 \dots i_m}| \leqslant M \sum_{j=1}^{m} ||x^{(j)}_{i_1 \dots i_m}|| \quad (11)$$

$$||x_{i_1 \dots i_m}|| \leqslant Md \Big(\sum_{t=1}^{r-1} \sum_{\Sigma k = t} \sum_{j=1}^{m} ||x^{(j)}_{k_1 \dots k_m}|| \cdot ||x_{i_1 - k_1 \dots i_m - k_m}|| + \sum_{j=1}^{m} ||x^{(j)}_{i_1 \dots i_m}|| \Big). \quad (12)$$

By adding all the inequalities (12) for $\Sigma i = r$ and denoting

$$\sum_{\Sigma i = r} ||x_{i_1 \dots i_m}|| = P_r, \quad P_0 = 1 \quad (13)$$

we obtain

$$P_r \leqslant Mmd \Big(\sum_{t=1}^{r-1} P_{t-1} P_{r-t} + P_{r-1} \Big) = c \sum_{t=1}^{r} P_{t-1} P_{r-t}$$

where $c = Mmd$. Defining a new sequence of constants Q_r by

$$Q_0 = 1, \quad Q_r = c \sum_{t=1}^{r} Q_{t-1} Q_{r-t}, \quad r = 1, 2, \dots$$

it is easy to show that

$$P_r \leqslant Q_r \quad (14)$$

and that the series

$$Q_0 + Q_1 z + Q_2 z^2 + \dots \quad (15)$$

is convergent for $|z| \leqslant R$ where $R = 1/(4c)$. In view of (11), (13), and (14) it is clear that

$$\sum_{\Sigma i = r} ||x_{i_1 \dots i_m}|| \leqslant Q_r, \quad \sum_{\Sigma i = r} |\lambda_{i_1 \dots i_m}| \leqslant MmQ_{r-1}$$

so that the series (15) dominates the series (4). Hence for all $|p_i| \leqslant R$ where

$$R = 1/(4mMd)$$

the series (4) are convergent. This proves the theorem.

By very well known theorems we can conclude that in the "cube" $|p_i| \leqslant R' < R$ the series (4) has derivatives of all orders with respect to parameters. Partial derivatives exist also in the case of multiple eigenvalues of A_0 but the functions need not to be differentiable at $p_1 = \dots = p_m = 0$.

Consider for instance a simple example of $n = 2$ with $\lambda_1 = \lambda_2$,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + y \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

The eigenvalues of A are $1 + x + 2y \pm \sqrt{(x^2 + y^2)}$. These are not differentiable at $x = y = 0$.

Subsequently we shall need the formulae for the first derivatives and a bound for the second derivatives of the eigenvalues of A at $p_1 = \dots = p_m = 0$.

From (10) with $r = 1$ we find

$$\Big(\frac{\partial \lambda_k}{\partial p_i} \Big)^{(0)} = x_k^T A_i x_k \quad (16)$$

and from (11) with $r = 2$ and (12) with $r = 1$ we obtain

$$\Big| \Big(\frac{\partial^2 \lambda_k}{\partial p_i \partial p_j} \Big)^{(0)} \Big| \leqslant 2M^2 d = L. \quad (17)$$

4. Numerical method

The analytic properties of the eigenvalues of A enable us to apply the general Newton's method for the solution of the system (2) with justification.

Given initial values of parameters $p_1^{(0)}, \dots, p_n^{(0)}$ and ordered eigenvalues $\lambda_1^* \geqslant \dots \geqslant \lambda_n^*$ we successively solve the systems of linear algebraic equations for $r = 0, 1, \dots$

$$\sum_{k=1}^{n} \lambda^{(r)}_{ik} \Delta p_k = \lambda_i^* - \lambda_i^{(r)} \quad (i = 1, \dots, n) \quad (18)$$

where

$$\lambda^{(r)}_{ik} = x_i^{(r)T} A_k x_i^{(r)} \quad (19)$$

from (16), and $\lambda_i^{(r)}, x_i^{(r)}$ form the eigensystem of the matrix

$$A^{(r)} = A_0 + \sum_{k=1}^{n} p_k^{(r)} A_k.$$

The calculated eigenvalues must also be ordered

$$\lambda_1^{(r)} \geqslant \ldots \geqslant \lambda_n^{(r)}. \tag{20}$$

Some kind of ordering of eigenvalues is essential since each right-hand side in (18) must approach zero.

The new approximation to parameters is then defined by

$$p_k^{(r+1)} = p_k^{(r)} + \Delta p_k \quad (k = 1, \ldots, n).$$

This process is repeated until the corrections to parameters are sufficiently small. If no convergence is observed after a certain number of steps a new initial approximation to parameters must be chosen.

We must remember that the formulae for partial derivatives (19) are correct only if in (20) no equality sign holds. It is possible to find the correct partial derivatives in the case of multiple eigenvalues too, but this would be useless because of the possible non-differentiability at such points. In the very unlikely case that at some intermediate stage some of the eigenvalues are exactly equal, we might do one step of iteration away from the solution we are approaching. But, if some of the calculated eigenvalues are only nearly equal we obtain from (19) perfectly correct derivatives. And, finally, if among given eigenvalues there are some multiple ones it causes no numerical difficulty.

5. Convergence of the method

One of the most useful convergence theorems for Newton's method is that of Kantorovich (1948) which applied to this problem combined with the results of Section 3 runs as follows:

Theorem. If all $\lambda_i^{(0)}$ are distinct, if the matrix $\Lambda^{(0)} = \{\lambda_{ik}^{(0)}\}$ is nonsingular, and if

$$h = N^2 L m n^2 \leqslant 1/2 \tag{21}$$

where $N = ||\Lambda^{(0)-1}||$, $m = \max_i |\lambda_i^* - \lambda_i^{(0)}|$, and L as defined in (17), then the process as described is quadratically convergent to a solution of the system (2).

This theorem is more practical than many others because all the quantities in (21) refer to the given initial point and are thus computable. Since in the neighbourhood of a solution N and L do not vary much the inequality (21) asserts a requirement on the quality of initial approximation. Needless to say, this requirement is, in general, too stringent and it is well known that in practice the process is convergent starting from much poorer initial approximations.

In our belief no simple bound can be obtained for N.

6. Numerical example

To illustrate the "normal" behaviour of the explained numerical process we shall consider an example of $n = 5$ with the following data:

$$A_0 = \begin{bmatrix} 0{\cdot}305 & & & & \\ 0{\cdot}672 & 0{\cdot}369 & & & \\ 0{\cdot}049 & 0{\cdot}231 & 0{\cdot}417 & & \\ 0{\cdot}806 & 0{\cdot}329 & 0{\cdot}801 & 0{\cdot}218 & \\ 0{\cdot}773 & 0{\cdot}215 & 0{\cdot}204 & 0{\cdot}914 & 0{\cdot}312 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 0{\cdot}870 & & & & \\ 0{\cdot}327 & 0{\cdot}953 & & & \\ 0{\cdot}145 & 0{\cdot}931 & 0{\cdot}743 & & \\ 0{\cdot}437 & 0{\cdot}201 & 0{\cdot}808 & 0{\cdot}986 & \\ 0{\cdot}451 & 0{\cdot}037 & 0{\cdot}300 & 0{\cdot}882 & 0{\cdot}926 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0{\cdot}733 & & & & \\ 0{\cdot}823 & 0{\cdot}142 & & & \\ 0{\cdot}109 & 0{\cdot}564 & 0{\cdot}768 & & \\ 0{\cdot}905 & 0{\cdot}414 & 0{\cdot}562 & 0{\cdot}736 & \\ 0{\cdot}114 & 0{\cdot}852 & 0{\cdot}639 & 0{\cdot}342 & 0{\cdot}788 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0{\cdot}212 & & & & \\ 0{\cdot}634 & 0{\cdot}891 & & & \\ 0{\cdot}421 & 0{\cdot}456 & 0{\cdot}720 & & \\ 0{\cdot}542 & 0{\cdot}577 & 0{\cdot}812 & 0{\cdot}345 & \\ 0{\cdot}882 & 0{\cdot}123 & 0{\cdot}420 & 0{\cdot}786 & 0{\cdot}743 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} 0{\cdot}341 & & & & \\ 0{\cdot}707 & 0{\cdot}452 & & & \\ 0{\cdot}921 & 0{\cdot}863 & 0{\cdot}990 & & \\ 0{\cdot}551 & 0{\cdot}010 & 0{\cdot}571 & 0{\cdot}412 & \\ 0{\cdot}822 & 0{\cdot}527 & 0{\cdot}517 & 0{\cdot}163 & 0{\cdot}234 \end{bmatrix}$$

$$A_5 = \begin{bmatrix} 0{\cdot}321 & & & & \\ 0{\cdot}207 & 0{\cdot}341 & & & \\ 0{\cdot}897 & 0{\cdot}342 & 0{\cdot}710 & & \\ 0{\cdot}911 & 0{\cdot}310 & 0{\cdot}631 & 0{\cdot}009 & \\ 0{\cdot}721 & 0{\cdot}213 & 0{\cdot}452 & 0{\cdot}780 & 0{\cdot}466 \end{bmatrix}$$

$\lambda_1^* = 4{\cdot}0216090$, $\lambda_2^* = 0{\cdot}61568326$, $\lambda_3^* = 0{\cdot}42495309$, $\lambda_4^* = -0{\cdot}65946669$, $\lambda_5^* = -1{\cdot}0619386$,
$p_1^{(0)} = p_2^{(0)} = p_3^{(0)} = p_4^{(0)} = p_5^{(0)} = 0$.
The symmetric matrices A_k are given by lower triangles only.

On the Elliott 503 computer of the University of Surrey and using the Francis QR method for diagonalization of symmetric tridiagonal matrices we have obtained the following iterates (**Table 1**).

It is interesting to follow the history of the eigenvalues which are given in **Table 2**.

The final eigenvalues agree with the given ones to 7 decimal places.

The exact result is:
$p_1^* = 0{\cdot}10, p_2^* = 0{\cdot}11, p_3^* = 0{\cdot}12, p_4^* = 0{\cdot}13, p_5^* = 0{\cdot}14$.

It may be verified that the condition for convergence is fulfilled only after 4th iteration at which stage a solution has been calculated correctly to 6 decimal places.

Table 1

r	$p_1^{(r)}$	$p_2^{(r)}$	$p_3^{(r)}$	$p_4^{(r)}$	$p_5^{(r)}$	$h^{(r)}$
0	0·00000000	0·00000000	0·00000000	0·00000000	0·00000000	206666·89
1	0·08268049	0·13503942	0·13597724	0·09493792	0·15998539	7502·8815
2	0·09923862	0·11076764	0·12183099	0·12872758	0·13931725	215·79632
3	0·09999730	0·11000218	0·12000549	0·12999819	0·13999653	0·66501625
4	0·10000008	0·10999995	0·11999990	0·13000032	0·13999975	0·02141011

Table 2

r	$\lambda_1^{(r)}$	$\lambda_2^{(r)}$	$\lambda_3^{(r)}$	$\lambda_4^{(r)}$	$\lambda_5^{(r)}$
0	2·4265380	0·50689792	0·21595877	−0·61679072	−0·91160397
1	4·0485474	0·61795635	0·40626636	−0·64914087	−1·0827891
2	4·0216960	0·61629394	0·42475279	−0·65958413	−1·0623187
3	4·0216093	0·61568511	0·42495200	−0·65946739	−1·0619393
4	4·0216090	0·61568327	0·42495310	−0·65946670	−1·0619385

7. Comments

We have had substantial numerical experience with this problem. In the majority of examples we have dealt with, the matrices were of orders up to 15, and with good initial approximations we have obtained results to 7 digit accuracy in about 6 iterations.

In some of the problems the number of parameters has been less than n. In such cases the least squares solution has been sought for; the behaviour of convergence has been much the same.

The number of possible solutions is theoretically $n!$ as the system is essentially an algebraic system of equations of orders $1, 2, \ldots, n$. None of the known methods for the solution of systems of non-linear equations provides us with all the solutions.

For the determination of other solutions we have had to rely on the rather *ad hoc* choice of different initial approximations.

It is possible to generalize the problem and the method in some ways. If the matrices (3) are non-symmetric the computation must be extended to the complex field, and in the calculation of derivatives one must employ the left-hand side eigenvectors too. If the matrix (1) is a non-linear function of parameters, the matrices A_k must be replaced by $\partial A/\partial p_k$ at a given point. We have had no numerical experience in such cases.

Acknowledgements

The work described here has been carried out during my research fellowship at the University of Surrey, London. I wish to thank Dr. J. H. Wilkinson of the National Physical Laboratory for careful reading of the manuscript and for making a number of valuable suggestions.

References

KANTOROVICH, L. V. (1948). Functional analysis and applied mathematics, *Uspehi Mat. Nauk*, Vol. 3, p. 89.
RELLICH, F. (1936). Störungstheorie der Spectralzerlegung I, *Math. Ann.*, Vol. 113, p. 600.
TOMAN, S., and PLIVA, J. (1966). The multiplicity of solutions of the inverse secular problem, *J. Mol. Spectry*, Vol. 21, p. 362.

Correspondence

To the Editor,
The Computer Journal.

Sir,

Orion FORTRAN compiler

Although one must support the praise for the Orion multiprogramming supervisor given by E. C. Willey ("An established U.K. software development", this *Journal*, Nov. 1967), one regrets that he did not mention another development, the Orion FORTRAN compiler (this *Journal*, July 1964). This compiler, produced by a joint project of the Rutherford Laboratory and the Orion manufacturers, was written in its own source language (which is close to Atlas FORTRAN), met its target dates, and initially implemented all specified language features. It is still used by one organization for all commercial processing, as well as for technical and scientific calculations, in preference to the Nebula system. Two errors were corrected after release for general use. It took about 10 man-years; how many did Nebula need?

Yours faithfully,

R. TAYLOR

Rutherford High Energy Laboratory,
Chilton,
Didcot,
Berks.

23 November 1967.