Numerical Method for Three-Parameter Eigenvalue Problems using Newton’s method based on Trace Theorem

Songita Boruah  
Research scholar  
Dept. Of Mathematics D.U  
Dibrugarh Assam, India

Arun Kumar Baruah  
Professor  
Dept. Of Mathematics D.U  
Dibrugarh Assam, India

ABSTRACT
In this paper Newton’s method using Trace Theorem for three-parameter eigenvalue problems are discussed and some numerical results are presented to illustrate the performance and application of the method.

General Terms
Eigenvalue, Eigenvector, Multiparameter

Key words
Multiparameter, eigenvalue, eigenvector, Newton’s method, Trace theorem

I. INTRODUCTION
Multiparameter eigenvalue problems are generalization of one-parameter eigenvalue problems and can be found when the method of separation of variables is applied to certain boundary value problems associated with partial differential equations. Typical examples are provided, for example by a vibrating membrane Roach [16] and a dynamical problem of homogeneous beam loaded by a vertical load Collatz [7].

Although the literature on multiparameter problems is over a century old, the attention of the mathematicians towards multiparameter problems started with the unifying work of Atkinson [1]. Much more works have been done in the field of one-parameter eigenvalue problems, both theoretically and numerically compared to two-parameter or more than two-parameter eigenvalue problems. Some works have been done theoretically in the field of multiparameter eigenvalue problems. Few authors namely Fox et. al [9], Sleeman [17], Baruah [3], Konwar [13], have dealt with the multiparameter eigenvalue problems numerically mainly in two parametric cases. Numerical methods applied to a three-parameter problems are very limited and hence some contribution in this area are always needed.

1.1 Three-Parameter Eigenvalue Problem And Its Reduction To A System Of One-Parameter Problems
A three-parameter eigenvalue problems in matrix form is as follows

\[ W_1(\lambda_x) = A_{10}x - \lambda_1 A_{11}x - \lambda_1 A_{12}x - \lambda_1 A_{13}x = 0 \]
\[ W_2(\lambda_y) = A_{20}y - \lambda_2 A_{21}y - \lambda_2 A_{22}y - \lambda_2 A_{23}y = 0 \]
\[ W_3(\lambda_z) = A_{30}z - \lambda_3 A_{31}z - \lambda_3 A_{32}z - \lambda_3 A_{33}z = 0 \]

Where \( \lambda_i \in \mathbb{R} \), \( i = 1,2,3 \) and

\[ x \in \mathbb{R} \setminus \{0\}, A_{10}, A_{11}, A_{12}, A_{13} \in \mathbb{R}^{n \times n} \]
\[ y \in \mathbb{R} \setminus \{0\}, A_{20}, A_{21}, A_{22}, A_{23} \in \mathbb{R}^{m \times m} \]
\[ z \in \mathbb{R} \setminus \{0\}, A_{30}, A_{31}, A_{32}, A_{33} \in \mathbb{R}^{p \times p} \]

Where \( \lambda_i \in \mathbb{R} \), \( i = 1,2,3 \) are the eigenvalues and \( x, y, z \) are called eigenvectors of the problem.

Problem (1.1.1) can be reduced to a system of three one-parameter problems:

\[ \Delta_1 u = \lambda_1 \Delta_1 u \]
\[ \Delta_2 u = \lambda_2 \Delta_2 u \]
\[ \Delta_3 u = \lambda_3 \Delta_3 u \] (1.1.2)

where \( \Delta_0, \Delta_1, \Delta_2, \Delta_3 \) are (mnp) × (mnp) dimensional matrices defined as

\[ \Delta_0 = A_{11} \otimes A_{22} \otimes A_{33} - A_{11} \otimes A_{23} \otimes A_{32} + A_{12} \otimes A_{23} \otimes A_{31} - A_{12} \otimes A_{21} \otimes A_{33} + A_{13} \otimes A_{21} \otimes A_{32} - A_{13} \otimes A_{22} \otimes A_{31} \] (1.1.3)
\[ \Delta_1 = A_{10} \otimes A_{22} \otimes A_{33} - A_{10} \otimes A_{23} \otimes A_{32} + A_{12} \otimes A_{23} \otimes A_{30} - A_{12} \otimes A_{20} \otimes A_{33} + A_{13} \otimes A_{20} \otimes A_{32} - A_{13} \otimes A_{22} \otimes A_{30} \] (1.1.4)
\[ \Delta_2 = A_{11} \otimes A_{20} \otimes A_{33} - A_{11} \otimes A_{23} \otimes A_{30} + A_{10} \otimes A_{23} \otimes A_{31} - A_{10} \otimes A_{21} \otimes A_{33} + A_{13} \otimes A_{21} \otimes A_{30} - A_{13} \otimes A_{20} \otimes A_{31} \] (1.1.5)
\[ \Delta_3 = A_{11} \otimes A_{20} \otimes A_{30} - A_{11} \otimes A_{20} \otimes A_{32} + A_{10} \otimes A_{20} \otimes A_{31} - A_{10} \otimes A_{21} \otimes A_{30} + A_{13} \otimes A_{21} \otimes A_{30} - A_{13} \otimes A_{20} \otimes A_{31} \] (1.1.6)

And

\[ u = x \otimes y \otimes z \]

With \( \otimes \) denoting the Kronecker product (or Tensor product) of two matrices discussed in (1.3).
Theorem: Let \((\lambda_1, \lambda_2, \lambda_3)\) be an eigenvalue and 
\((x, y, z)\) a corresponding eigenvector of the system 
\((1.1.1)\) then \((\lambda_1, \lambda_2, \lambda_3)\) is an eigenvalue of the system 
\((1.1.2)\) and \(u = x \otimes y \otimes z\) is the corresponding 
eigenvector.

1.2 The Kronecker Product
Definition 1.2.1: The Kronecker product 
\((\otimes, \otimes) : \mathbb{F}^{m \times n} \times \mathbb{F}^{p \times q} \to \mathbb{F}^{mn \times pq}\) is defined by

\[
A \otimes B = \begin{bmatrix}
    a_{11}B & L & a_{1n}B \\
     M & O & M \\
    a_{n1}B & L & a_{nn}B
\end{bmatrix}
\]

Where we use the standard notation \(A_{ij} = a_{ij}\)

The Kronecker product is a special case of the tensor product, and as such it inherits the properties of bilinearity and associativity, i.e.

\[
(kA) \otimes (kB) = K(A \otimes B) \\
A \otimes (B + C) = A \otimes B + A \otimes C \\
(A + B) \otimes C = (A \otimes C) + B \otimes C
\]

We now establish a famous property of the Kronecker product, from [9].

Lemma (Mixed product property). Let 
\(A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{p \times q}, C \in \mathbb{F}^{n \times k}, D \in \mathbb{F}^{q \times r}\). Then 
\((A \otimes B)(C \otimes D) = (AC \otimes BD)\). In particular, if \(A, B \in \mathbb{F}^{m \times n}\) and \(x, y \in \mathbb{F}^m\) then 
\((A \otimes B)(x \otimes y) = Ax \otimes By\).

2. NEWTON METHOD USIN TRACE THEOREM
Let \(f(\lambda) = \det W_1(\lambda) - f_1(\lambda)\), \(f_2(\lambda)\) and \(f_3(\lambda)\) (2.1)

\[
\begin{align*}
\frac{d}{dc}(\det B(c)) & = \det B(c) \cdot tr[B^{-1}(c) \cdot \frac{dB(c)}{dc}]
\end{align*}
\]

Using trace theorem, we have

\[
\frac{\partial f(\lambda)}{\partial \lambda_j} = \frac{\partial (\det W_1(\lambda))}{\partial \lambda_j} = tr[adj(W_1(c)) \cdot \frac{\partial W_1(c)}{\partial \lambda_j}]
\]

Then the 3 \times 3 Jacobian \(J_f(\lambda)\) of the function \(f(\lambda)\) is

\[
J_f(\lambda) = \text{tr}[adj(W_1(c)) \cdot \frac{\partial W_1(c)}{\partial \lambda_j}]
\]

If \(\det W_1(c) \neq 0\), then

\[
J_f(\lambda) = \det W_1(c) \cdot \frac{\partial W_1(c)}{\partial \lambda_j}
\]

Newton’s method to approach a solution of (2.2) has the form

\[
\lambda^{(n+1)} = \lambda^{(n)} - J_f^{-1}(\lambda^{(n)}) f(\lambda^{(n)}), n = 0, 1, 2, \ldots
\]

Where \(\lambda^{(0)} = (\lambda_{1}^{(0)}, \lambda_{2}^{(0)}, \lambda_{3}^{(0)})^T\) is an initial approximation.

3. NUMERICAL EXAMPLE:
We now present a numerical example to show the behaviour and application of our method

Consider the following system of equations

\[
\begin{pmatrix}
    1 & 0 \\
    0 & 2
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2
\end{pmatrix}
= \lambda_1
\begin{pmatrix}
    3 & 0 \\
    0 & 4
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2
\end{pmatrix}
+ \lambda_2
\begin{pmatrix}
    5 & 0 \\
    0 & 6
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2
\end{pmatrix}
\]

\[
\begin{pmatrix}
    7 & 0 \\
    0 & 8
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2
\end{pmatrix}
\]

To estimate rough bounds of the parameters \(\lambda_1, \lambda_2, \lambda_3\) of 
(3.1.1),(3.1.2),(3.1.3) we go through Gerschgorin theorem.
Theorem: Let $D_i$ be a circle whose centre is $a_{ii}$ and whose radius is $\sum_{j\neq i} |a_{ij}|$, where $j=1,2,3,\ldots,n$ and $j \neq i$. Then Gerschgorin says that

1. Every eigenvalue of $A$ must lie in the union of those circles.
2. If $k$ of these circles do not touch the other $n-k$ circles, then exactly $k$ eigenvalues lie in the union of those $k$ circles.

Using (1.2), we can write (3.1), (3.2), (3.3) as

$$\text{Diag}(0.0,0.36,-432,-72,-108,-36)u = \lambda_i u$$
$$\text{Diag}(480,84,600,144,264,40,582,152)u$$

Or,

$$\text{Diag}(0.0,0.2500,1.6364,-1.8000,-1.8564,-2368)u = \lambda_1 u$$
$$\text{Diag}(-72.0,-90.0,-72.24,-98.48)u = \lambda_2 u$$
$$\text{Diag}(480,84,600,144,264,40,582,152)u$$

Or,

$$\text{Diag}(-1500.00,-1500.00,-2727.00,-6000.00,-3402.00,-3158)u = \lambda_2 u$$
$$\text{Diag}(120.12,150.36,336.64,348.92)u = \lambda_1 u$$
$$\text{Diag}(480,84,600,144,264,40,582,152)u$$

Or,

$$\text{Diag}(2500.00,1429.00,2500.00,12727.16000.00,5979.00,6053)u = \lambda_2 u$$
$$\text{Diag}(1.6364,0.2727,1.2727)$$

Applying Gerschgorin’s Theorem we have

$$|\lambda_1| \leq 0.0559, 0.3402, 0.5979$$
$$|\lambda_2| \leq 0.1410, 0.3102, 0.2592$$
$$|\lambda_3| \leq 0.1845, 0.3402, 0.2500$$

Using these rough bounds of eigenvalues we consider a variety of starting values to calculate the approximate eigenvalues. Eight experiments are displayed in Table 1 and Table 2.

### Table 1

<table>
<thead>
<tr>
<th>Starting values</th>
<th>Iteration</th>
<th>$\lambda^{(k)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda^{(0)} = \begin{pmatrix} -1 \ -0.15 \ -0.3 \end{pmatrix}$</td>
<td>0</td>
<td>$(-1,-0.15,.3)^T$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$(-0.329,-1.441,2.597)^T$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$(-0.167,-1.473,2.552)^T$</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>$(-0.004,-1.499,2.501)^T$</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>$(-0.002,-1.500,2.501)^T$</td>
</tr>
<tr>
<td>$\lambda^{(0)} = \begin{pmatrix} -0.5 \ -0.1 \ -0.2 \end{pmatrix}$</td>
<td>0</td>
<td>$(-0.5,-0.1,2)^T$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$(0.0559,0.0606,0.0597)^T$</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>$(0.0001,0.0002,1.427)^T$</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>$\lambda^{(0)}$</th>
<th>$\lambda^{(4)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(0.0000,0.0001,0.1429)^T$</td>
<td>$(-1,0.519)^T$</td>
</tr>
<tr>
<td>$(-2.2201,0.519)^T$</td>
<td>$(-1.2201,0.519)^T$</td>
</tr>
<tr>
<td>$(-1.7635,0.519)^T$</td>
<td>$(-1.7635,0.519)^T$</td>
</tr>
<tr>
<td>$(-1.6364,0.2727)^T$</td>
<td>$(-1.6364,0.2727)^T$</td>
</tr>
<tr>
<td>$(-3.3102,0.2517)^T$</td>
<td>$(-3.3102,0.2517)^T$</td>
</tr>
<tr>
<td>$(-2.592,0.2503)^T$</td>
<td>$(-2.592,0.2503)^T$</td>
</tr>
<tr>
<td>$(-2.503,0.2500)^T$</td>
<td>$(-2.503,0.2500)^T$</td>
</tr>
<tr>
<td>$(-2.500,0.2500)^T$</td>
<td>$(-2.500,0.2500)^T$</td>
</tr>
<tr>
<td>$(-3.34,0.32,0.591)^T$</td>
<td>$(-3.34,0.32,0.591)^T$</td>
</tr>
<tr>
<td>$(-1.401,0.584)^T$</td>
<td>$(-1.401,0.584)^T$</td>
</tr>
<tr>
<td>$(-1.821,0.5967)^T$</td>
<td>$(-1.821,0.5967)^T$</td>
</tr>
<tr>
<td>$(-1.855,0.5979)^T$</td>
<td>$(-1.855,0.5979)^T$</td>
</tr>
<tr>
<td>$(-1.856,0.5979)^T$</td>
<td>$(-1.856,0.5979)^T$</td>
</tr>
<tr>
<td>$(-2.1,-0.75,1.8)^T$</td>
<td>$(-2.1,-0.75,1.8)^T$</td>
</tr>
<tr>
<td>$(-1.8027,0.6011,0.6024)^T$</td>
<td>$(-1.8027,0.6011,0.6024)^T$</td>
</tr>
<tr>
<td>$(-1.8000,0.6000,0.6000)^T$</td>
<td>$(-1.8000,0.6000,0.6000)^T$</td>
</tr>
<tr>
<td>$(-0.5,-0.5,1)^T$</td>
<td>$(-0.5,-0.5,1)^T$</td>
</tr>
<tr>
<td>$(0.3276,0.3793,0.7414)^T$</td>
<td>$(0.3276,0.3793,0.7414)^T$</td>
</tr>
<tr>
<td>$(-2.566,0.6348)^T$</td>
<td>$(-2.566,0.6348)^T$</td>
</tr>
<tr>
<td>$(-2.382,0.6074)^T$</td>
<td>$(-2.382,0.6074)^T$</td>
</tr>
<tr>
<td>$(-2.369,0.3158,0.6053)^T$</td>
<td>$(-2.369,0.3158,0.6053)^T$</td>
</tr>
<tr>
<td>$(-2.368,0.3158,0.6053)^T$</td>
<td>$(-2.368,0.3158,0.6053)^T$</td>
</tr>
</tbody>
</table>
The approximate eigenvalues obtained from Table 1 are 
\((-0.004, -1.499, 2.501)^T, (0.0, 1.429)^T, (-1.6364, -2.727, 1.2727)^T, 
(-25.0, 25)^T, (-1.185, -3.402, 5.979)^T, (-1.8, -6.1.6)^T, 
(-2368, -3.158, 6.053)^T\)

The successive difference between the eigenvalues obtained from table1 and the successive values of \(f(\lambda)\) are calculated in Table2.

### 4. CONCLUSION

Table 2 shows that the successive difference between the eigenvalues are gradually decreases and the values of \(f(\lambda)\) also ceases to zero. So the method converges to the exact eigenvalues rapidly. From Table 2 we can say that this method is computationally attractive and we can use this method easily to solve three-parameter eigenvalue problems. In the process describe above the approximate eigenvalues can be obtained more easily and the convergence is almost guaranteed as well as being more rapid. It will definitely play significant roles for further research in tackling the three-parameter for matrices and the multiparameter problem in general.

### 5. REFERENCES


---

**Table 2**

<table>
<thead>
<tr>
<th>Starting values (\lambda^{(0)})</th>
<th>Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>[\begin{pmatrix} -1 \ -15 \ .3 \end{pmatrix}]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td>[\begin{pmatrix} .1 \ -.1 \ .2 \end{pmatrix}]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>[\begin{pmatrix} -1 \ -1 \ 1 \end{pmatrix}]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>[\begin{pmatrix} -.3 \ .1 \ .2 \end{pmatrix}]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>[\begin{pmatrix} -.34 \ -.32 \ .591 \end{pmatrix}]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>[\begin{pmatrix} -2.1 \ -.75 \ 1.8 \end{pmatrix}]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>[\begin{pmatrix} -.5 \ -.5 \ 1 \end{pmatrix}]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>


