Mysteries of Eigenvalues, Eigenvectors & their Applications in the Diagonalization of a Matrix & in the Cayley-Hamilton Theorem to Find the Matrix Inverse

Kamalmani Baral
Pulchowk Campus, Institute of Engineering Tribhuvan University
Corresponding email: kamalmanib@yahoo.com

Abstract: In this paper we deal with eigenvalues and eigenvectors (E-values & E-vectors) in diagonalizing a square matrix and in the Cayley-Hamilton theorem used to find the inverse of a given square matrix.

1. Introduction

If \( A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2j} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nj} & \ldots & a_{nn} \end{bmatrix} \) is a square matrix of order \( n \) then the matrix

\[ (A - \lambda I) \]

is called the characteristic matrix of \( A \). If the determinant of this characteristic matrix of \( A \) is taken and equated to zero i.e.

\[ |A - \lambda I| = 0 \]  \hspace{1cm} (i)

i.e.
\[ a_{11} - \lambda \quad a_{12} \quad \ldots \quad a_{1n} \\
\vdots \quad \vdots \quad \ldots \quad \vdots \quad \vdots \\
a_{nj} \quad a_{n2} \quad \ldots \quad a_{nn} \]

\[ = 0 \]

\[ \lambda^n + p_0 \lambda^{n-1} + \ldots + p_n = 0 \] (ii)

it is called the characteristic equation of A. This equation naturally has n roots, so these n-roots of \( \lambda \) are called the characteristic roots, latent roots, or eigenvalues of A.

Again, if \( A = [a_{ij}]_{n \times n} \) is a square matrix of order n and \( X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \) is a column vector which is transformed by A into its scalar multiple such that

\[ AX = \lambda X \] (iii)

And if I is the unit matrix of the same order then (iii) can be written as

\[ AX = \lambda IX \]

or

\[ (A - \lambda I) X = 0 \] (iv)

which is called the characteristic matrix equation and this equation (iv) represents n homogeneous equations:

\[ (a_{11} - \lambda) x_1 + a_{12} x_2 + \ldots + a_{1n} x_n = 0 \]
\[ a_{21} x_1 + (a_{22} - \lambda) x_2 + \ldots + a_{2n} x_n = 0 \]
\[ \vdots \quad \vdots \quad \ldots \quad \vdots \quad \ldots \quad \ldots \]
\[ a_{n1} x_1 + a_{n2} x_2 + \ldots + (a_{nn} - \lambda) x_n = 0 \]

This system of n homogeneous linear equations will have a non-zero solution if \( |A - \lambda I| \) is singular i.e. \( |A - \lambda I| = 0 \). Moreover the roots of \( (A - \lambda I) = 0 \) gives the n-eigenvalues \( \lambda_1, \ldots, \lambda_n \). To each eigenvalue of A, there corresponds a non-zero solution to the vector

\[ X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \]

This vector X is called the eigenvector of A (corresponding to that particular eigenvalue of A).

Note that the eigenvector corresponding to each eigenvalue is not unique. So any scalar multiple of the E-vector are also the E-vectors of A corresponding to those particular E-values of A.
For example, if \( A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \) is a square matrix then the characteristic equation is

\[
|A - \lambda I| = 0
\]

i.e.

\[
\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = 0 \Rightarrow (\lambda^2 - 4\lambda + 3) = 0 \Rightarrow \lambda = 1 \text{ or } 3
\]

which are the E-values of \( A \). Again, if \( \begin{bmatrix} x \\ y \end{bmatrix} \) is an E-vector then the corresponding matrix equation is

\[
(A - \lambda I) x = 0 \quad (v)
\]

i.e.,

\[
\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

\[(vi)\]

When \( \lambda = 1 \) then (vi) reduces to

\[
\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} x + y = 0 \\ x + y = 0 \end{cases}
\]

(No of LI solutions = unknowns-rank = 2-1=1)

If we assign \( x = 1 \) then we get \( y = -1 \) so the E-vector is \( \begin{bmatrix} 1 \\ -1 \end{bmatrix} \) and the set of all E-vectors is

\[
\begin{bmatrix} k_1 \\ -k_1 \end{bmatrix} (k_1 \neq 0). \text{ Similarly when } \lambda = 3 \text{ then (vi) give the E-vector } \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and the set of all eigenvectors is } \begin{bmatrix} k_2 \\ k_2 \end{bmatrix} (k_2 \neq 0).\]

2. **Application of E-values and E-vectors in the Diagonalization of a Square Matrix**
Now we shed light on the reduction of a given square matrix into the diagonal form through the transforming matrix $P$ to the diagonal form $D$.

It is clear that "If a square matrix $A = [a_{ij}]_{n \times n}$ of order $n$ has $n$ linearly independent E-vectors then a matrix $P$ can be found such that $P^{-1}AP$ is a diagonal matrix (whose diagonal elements are the same E-values)". For example, if we are given a matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

then a transforming matrix $P$ can be found such that $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ (denoted as $D$)

which transforms $A$ into the diagonal form $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ whose diagonal elements are the same E-values.

(Note that the sum of all diagonal elements of a given square matrix $A$ is equal to the sum of all E-values of $A$).

Moreover, if $P = [X_1, X_2]$ be a column vector of the transforming matrix $P$ of a given square matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. The E-values of $A$ are 1 and 3 and the E-vectors corresponding to $\lambda=1$ and $\lambda=3$ are $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ ($=X_1$) and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ($=X_2$) respectively.

Then we make a new matrix $P = [X_1, X_2] = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

which is the transforming matrix which transforms $A$ into the diagonal from $D$ i.e.

$$P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

which is the diagonal from $D$ and the diagonal elements are the same E-values. This is the direct relation among the E-values, E-vectors, transforming matrix $P$ and the diagonal form $D$ of a given square matrix $A$ which has a relation like a grandfather.

This relationship can be sketched/represented graphically as shown below.
This pictorial representation shows the deep relationship, among the E-values and the E-vectors of a given square matrix $A$.

### 3. Application of Cayley-Hamilton theorem in finding the inverse of a given square matrix

Here we show the application of the Cayley-Hamilton Theorem (C-H theorem) to find the inverse of a given square matrix in simpler form. We know that the Cayley-Hamilton theorem states that "every square matrix $A$ satisfies its own characteristic equation".

For, if $A = [a_{ij}]_{n \times n}$ is a square matrix of order $n$ then its characteristic equation is

$$|A - \lambda I| = 0$$

i.e. $\lambda^n + p_1 \lambda^{n-1} + p_2 \lambda^{n-2} + \ldots + p_n I = 0$ \hspace{1cm} (vii)

then by C-H theorem $A^n + p_1 A^{n-1} + p_2 A^{n-2} + \ldots + p_n I = 0$ \hspace{1cm} (viii)

where $I$ is the unit matrix of the same order. The inverse of the given square matrix $A$ can be obtained directly by using this Cayley-Hamilton theorem. From the equation (viii)

$$A^n + p_1 A^{n-1} + p_2 A^{n-2} + \ldots + p_n I = 0$$

Now multiplying both sides by $A^{-1}$

we get $A^{-1} (A^n + p_1 A^{n-1} + \ldots + p_n I) = A^{-1}.0$

or $A^{n-1} + p_1 A^{n-2} + \ldots + p_n I + A^{-1}. p_n = 0$

$$\Rightarrow A^{-1} = \frac{1}{p_n} \left[ A^{n-1} + p_1 A^{n-2} + \ldots + p_{n-1} I \right]$$

results in the inverse of a square matrix $A$. This is a very nice application of the Cayley-Hamilton theorem from which we get $A^{-1}$ by a short and sweet method without using the lengthy formula

$$A^{-1} = \frac{\text{Adj} \ A}{|A|} \quad (A \neq 0)$$

which are nice applications of E-values & E-vectors in the diagonalization of a matrix & in the Cayley-Hamilton theorem.

**Some common properties of E-values, E-vectors, transforming matrix & diagonal form (matrix)**

1. Any square matrix $A$ and its transpose have the same E-values.
2. The trace of the matrix equals to the sum of E-values of the same matrix.
   i.e. $(a_{11} + a_{22} + \ldots + a_{nn} = \lambda_1 + \lambda_2 + \ldots + \lambda_n)$
3. The determinant of the matrix $A$ equals the product of E-values of $A$.
4. If $(\lambda_1, \lambda_2 \ldots \lambda_n)$ are the $n$ E-values of $A$ then the E-values of $kA$ are $(k\lambda_1, k\lambda_2, \ldots k\lambda_n)$.
5. E-values of $A^{-1}$ are $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots \frac{1}{\lambda_n}$.
6. If $\lambda$ is an E-value of a matrix $A$ then $\frac{1}{\lambda}$ is an E-value of $A^{-1}$.
7. If $(\lambda_1, \lambda_2 \ldots \lambda_n)$ are the E-values of a square matrix $A$, then $A^m$ has the E-values $\lambda_1^m, \lambda_2^m, \ldots, \lambda_n^m$, where $m$ is a positive integer.
8. Zero is an E-value of a square matrix $A$ if and only if $A$ is singular i.e. $|A| = 0$.
9. The E-values of a triangular matrix are just the diagonals elements of the matrix.
10. Two matrices $A$ & $P^{-1}$ $AP$ have the same E-values.

### Eigen values of some special type of matrices

1. The E-values of the Hermitian matrix are real.
2. The E-values of a real symmetric matrix are all real.
3. Every E-value of a skew-Hermitian matrix is either 0 or a pure imaginary number.
4. The E-values of a unitary matrix are of unit modulus.
5. The E-values of an orthogonal matrix are of unit modulus.
6. If $\lambda$ is an E-value of an orthogonal matrix $A$ then $\frac{1}{\lambda}$ is also an E-value of $A$.

### 4. Conclusion

E-values, E-vectors have a nice application in the diagonalization of a square matrix and in the Cayley-Hamilton theorem used to find the matrix inverse and in all the endeavors of medical, engineering and social sciences.

### Reference

[1] Elementary Linear Algebra, Keith Matthews, Lecture notes and solutions from 1991 in PDF or post script