Presentation 3: Eigenvalues and Eigenvectors of a Matrix

Order of Presentation:
1. Definitions of Eigenvalues and Eigenvectors
2. Mathematical Expression of Eigenvectors and Eigenvalues
3. Computing Eigenvalues and Eigenvectors
   • Characteristic Equation
4. Properties of Eigenvalues and Eigenvectors
5. Linear Independence of Eigenvectors
6. Geometric Interpretation
7. Triangular Matrices
8. Singular Matrices
9. Example for 2 x 2 matrices
10. Applications of Eigenvalues and Eigenvectors

1. Definitions:
   • Eigenvector: An eigenvector of a square matrix is a non-zero vector, that after being multiplied by the matrix, remains parallel to the original vector.
   • Eigenvalue: The eigenvalue is the value that the eigenvector is scaled by when multiplied by the matrix.
   • Eigenspace: The eigenspace is the set of all eigenvectors that have the same eigenvalue, and it also includes the zero vector.

2. Mathematical Expression:
   \[ A \cdot v = \lambda \cdot v \]
   • A is the square matrix.
   • v is the non-zero eigenvector of A.
   • \lambda is the eigenvalue of A that corresponds to v. It is the scalar that when multiplied by the eigenvector, produces \( A \cdot v \).
   • These properties exist for every size square matrix. A 3x3 example is:
   \[
   \begin{pmatrix}
   a_1 & b_1 & c_1 \\
   a_2 & b_2 & c_2 \\
   a_3 & b_3 & c_3
   \end{pmatrix}
   \begin{pmatrix}
   x \\
   y \\
   z
   \end{pmatrix}
   = \lambda
   \begin{pmatrix}
   x \\
   y \\
   z
   \end{pmatrix}
   \]

3. Computing Eigenvalues and Eigenvectors:
   \[(A - \lambda I) v = 0\]
   • This is the characteristic equation.
   • A is the square matrix; v is the eigenvector, and \lambda is the eigenvalue just as in the original equation.
   • I is the n x n identity matrix.
   • \(A - \lambda I\) cannot be invertible.
   • The equation is solved for v to find the eigenvector.
   • The eigenvalues are found by finding the roots of the equation below:
   \[\det (A - \lambda I) = 0\]
4. Properties:
   1. If $A$ is a triangular matrix, then the eigenvalues of $A$ are the entries on its main diagonal.
   2. 0 is an eigenvalue of $A$ if and only if $A$ is not invertible. The equation $(A - \lambda I) v = 0$ has a nontrivial solution if and only if $A$ is not invertible.
   3. If $v_1, \ldots, v_r$ are the eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix $A$, then the set \{v_1, \ldots, v_r\} is linearly independent.

5. Linear Independence:
   If $v_1, \ldots, v_r$ are the eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix $A$, then the set \{v_1, \ldots, v_r\} is linearly independent.

Proof:
If $v_1, \ldots, v_r$ is linearly dependent, then there is a least index $p$ such that $v_{p+1}$ is a linear combination of the preceding (linearly independent) vectors, and there exist scalars $c_1, \ldots, c_p$ such that
   \[ c_1 v_1 + \ldots + c_p v_p = v_{p+1} \]  
(1)

Multiplying both sides of (1) by $A$ and using the fact that $A v_k = \lambda_k v_k$ for each $k$:
   \[ c_1 A v_1 + \ldots + c_p A v_p = A v_{p+1} \]
   \[ c_1 \lambda_1 v_1 + \ldots + c_p \lambda_p v_p = \lambda_{p+1} v_{p+1} \]  
(2)

Multiplying both sides of (1) by $\lambda_{p+1}$ and subtracting the result from (2):
   \[ c_1 (\lambda_1 - \lambda_{p+1}) v_1 + \ldots + c_p (\lambda_p - \lambda_{p+1}) v_p = 0 \]  
(3)

Since \{v_1, \ldots, v_r\} is linearly independent, the weights in (3) are all zero. But none of the factors $\lambda_i - \lambda_{p+1}$ are zero, because the eigenvalues are distinct. Hence $c_i = 0$ for $i = 1, \ldots, p$. But then (1) says that $v_{p+1} = 0$, which is impossible. Hence \{v_1, \ldots, v_r\} cannot be linearly dependent and therefore must be linearly independent.

6. Geometric Interpretation of Eigenvalues and Eigenvectors:
   The eigenvector $v$ is scaled by the eigenvalue $\lambda$.
   \[ A \hat{v} = \lambda \hat{v} \quad (\lambda > 1) \]
   \[ A \hat{v} = \lambda \hat{v} \quad (0 < \lambda < 1) \]
   \[ A \hat{v} = \lambda \hat{v} \quad (\lambda < 0) \]
Examples:
- These examples contain the same matrix $A$, but use different eigenvectors and eigenvalues.

When $\lambda = 2$:
\[
\begin{bmatrix}
2 & 1 \\
0 & -1
\end{bmatrix}
\begin{bmatrix}
1 \\
0
\end{bmatrix}
= 
\begin{bmatrix}
2 \\
0
\end{bmatrix}
\]

When $\lambda = -1$
\[
\begin{bmatrix}
2 & 1 \\
0 & -1
\end{bmatrix}
\begin{bmatrix}
1 \\
3
\end{bmatrix}
= 
\begin{bmatrix}
-1 \\
3
\end{bmatrix}
\]

7. Triangular Matrices:

For 2x2 matrices:
\[
\begin{bmatrix}
0 & 1 \\
0 & 0
\end{bmatrix}
\]

For 3x3 matrices:
\[
\begin{bmatrix}
\alpha_{11} & 0 & 0 \\
\alpha_{21} & \alpha_{22} & 0 \\
\alpha_{31} & \alpha_{32} & \alpha_{33}
\end{bmatrix}
\]

- The eigenvalues of a triangular matrix are those in the main diagonal of the matrix.
- In the matrix shown above, the eigenvalues would be $\alpha_{11}$, $\alpha_{22}$, and $\alpha_{33}$.
- The identity matrix is an example of a triangular matrix.
- There exist triangular matrices for square matrices of any size.

8. Singular Matrices:

- A singular matrix is a square matrix that does not have a matrix inverse.
- A matrix is singular if and only if its determinant is 0
  - Matrix $A$ is singular $\iff$ $\det (A) = 0$.
- When using the characteristic equation: $(A - \lambda I) v = 0$, the matrix $(A - \lambda I)$ must be a singular matrix because $v$ is a non-zero vector.
9. Example for 2x2 Matrices:

Finding Eigenvalues:

\[ A = \begin{bmatrix} 6 & 16 \\ -1 & -4 \end{bmatrix} \quad \lambda I - A = \begin{bmatrix} \lambda - 6 & -16 \\ 1 & \lambda + 4 \end{bmatrix} \]

Characteristic Polynomial: \[ p(\lambda) = (\lambda - 6)(\lambda + 4) + 16 = \lambda^2 - 2\lambda - 8 = 0 \]

So the eigenvalues are \( \lambda = 4 \) and \( \lambda = -2 \)

Finding Eigenvectors:

• When \( \lambda = 4 \):

\[ A \begin{bmatrix} x \\ y \end{bmatrix} = 4 \begin{bmatrix} x \\ y \end{bmatrix} \]

\[ \begin{bmatrix} 6 & 16 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 4 \begin{bmatrix} x \\ y \end{bmatrix} \]

Solving: \[ 6x + 16y = 4x \quad \rightarrow \quad 2x + 16y = 0 \]
\[ -x - 4y = 4y \quad \rightarrow \quad x + 8y = 0 \]
\[ y = 1, \quad x = -8 \]

Eigenvector:
\[ t \begin{bmatrix} -8 \\ 1 \end{bmatrix} \]

• When \( \lambda = -2 \):

\[ \begin{bmatrix} 6 & 16 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -2 \begin{bmatrix} x \\ y \end{bmatrix} \]

Solving: \[ 6x + 16y = -2x \quad \rightarrow \quad 8x + 16y = 0 \]
\[ -x - 4y = -2y \quad \rightarrow \quad x + 2y = 0 \]
\[ y = 1, \quad x = -2 \]

Eigenvector:
\[ t \begin{bmatrix} -2 \\ 1 \end{bmatrix} \]

10. Applications of Eigenvalues and Eigenvectors:

Markov Chains:

• Definitions and Remarks:

  Transition Matrix: matrix in which the entries represent the probability of transition from the state corresponding to \( i \) to the state corresponding to \( j \)

  If all of the entries in a transition matrix \( T \) are between 0 and 1, then the matrix \( T^k \) will converge as \( k \rightarrow \infty \).

  If the matrix has entries of 0 and 1, \( T^k \) may still converge, or it may oscillate.
Markov Chain Example: Delivery drivers for a car rental agency

Transition Matrix:

\[ T = \begin{bmatrix}
A & B & C \\
0.3 & 0.4 & 0.5 \\
0.3 & 0.4 & 0.3 \\
0.4 & 0.2 & 0.2
\end{bmatrix} \]

Where each column represents the starting location and each row represents the ending location.

Two Deliveries:
- If starting in location C, to find the probability of being in location B after two deliveries:
  \[ P(CA)P(AB) + P(CB)P(BB) + P(CC)P(CB) = (0.5)(0.3) + (0.3)(0.4) + (0.2)(0.3) = 0.33 \]
  This means that there is a 33% chance of being in location B after two deliveries if starting in location C. This value is the same as taking the inner product of row 2 and column 3.

Three Deliveries:
- If starting in location C, to find the probability of being in location B after three deliveries:
  We must use the matrix \( T^2 \) as well as the original transition matrix \( T \).
  \[ T^2 = \begin{bmatrix}
A & B & C \\
0.41 & 0.38 & 0.37 \\
0.33 & 0.34 & 0.33 \\
0.26 & 0.28 & 0.27
\end{bmatrix} \]
  The probability will be the inner product of row 2 of \( T^2 \) and column 3 of \( T \):
  \[ p(CA)P(AB) + p(CB)P(BB) + p(CC)P(CB) = (0.37)(0.3) + (0.33)(0.4) + (0.3)(0.3) = 0.333 \]
  As \( k \) approaches \( \infty \), \( T^k \) will converge to:
  \[ \begin{bmatrix}
0.38 & 0.38 & 0.38 \\
0.33 & 0.33 & 0.33 \\
0.27 & 0.27 & 0.27
\end{bmatrix} \]
  Note that each column has the same entries.

If there are 54 delivery people, and they are divided up evenly among the locations A, B, and C, then after one delivery:
\[
\begin{bmatrix}
0.3 & 0.4 & 0.5 \\
0.3 & 0.4 & 0.3 \\
0.4 & 0.2 & 0.2
\end{bmatrix} \begin{bmatrix}
18 \\
18 \\
18
\end{bmatrix} = \begin{bmatrix}
21.6 \\
18 \\
14.4
\end{bmatrix}
\]
  This only gives an approximation of how many delivery people will be at each location after one delivery (because of course you will not have fractions of people). The fractions are due to the expected values.

After many deliveries:
\[
\begin{bmatrix}
0.38 & 0.38 & 0.38 \\
0.33 & 0.33 & 0.33 \\
0.27 & 0.27 & 0.27
\end{bmatrix} \begin{bmatrix}
18 \\
18 \\
18
\end{bmatrix} = \begin{bmatrix}
21 \\
18 \\
15
\end{bmatrix}
\]
This vector is the same as the vector produced by multiplying 54 by any column of the matrix $T^k$ as $k \to \infty$:

$$
\begin{bmatrix}
0.38 \\
0.33 \\
0.27
\end{bmatrix}
= 
\begin{bmatrix}
21 \\
18 \\
15
\end{bmatrix}
$$

Markov Chain Theory:

- A transition matrix is **regular** if all of the entries of $T^k$ for some integer $k$ are positive.
- The Markov chain process is regular if its transition matrix is regular.
- The **state vector** is a column vector whose $i$th component represents the probability that the system is in the $i$th state at that time.
- If $T$ is a regular transition matrix, then as $k$ approaches infinity, $T^k \to S$ where $S$ is a matrix of the form $[v, v, \ldots, v]$ with $v$ being a constant vector.
- If $T$ is a regular transition matrix of a Markov chain process, and if $X$ is any state vector, then as $k$ approaches infinity, $T^k X \to p$, where $p$ is a fixed probability vector (the sum of its entries is 1), all of whose entries are positive.
- The **steady-state vector** of a regular Markov chain with transition matrix $T$ is the unique probability vector $p$ satisfying $Tp=p$.
- Another way to consider this is:
  If a Markov chain with a regular transition matrix $T$ has $M$ as the limit of $T^k$ as $k \to \infty$, then $T^k X \to MX = p$. ($X$ is any state vector.) Therefore, the system approaches the steady-state vector $p$.
- To compute the steady-state vector of a regular Markov chain, solve $Tp=p$, for $p$.

Relationship to Eigenvectors and Eigenvalues:

- A steady-state vector of a regular Markov chain is an eigenvector for the transition matrix corresponding to the eigenvalue 1.
- The eigenvalues of the matrix $T$ are the solutions to the equation $\det(T-\lambda I)=0$ where $I$ is the identity matrix of the same size as $T$.
- If $\lambda$ is an eigenvalue of $T$, then an eigenvector corresponding to $\lambda$ is a non-zero solution to the homogeneous system $(T-\lambda I)X=0$.
- There are infinitely many eigenvectors corresponding to a fixed eigenvalue.

Back to Example:

$$
TX = \lambda X
$$

$$
\begin{bmatrix}
0.3 & 0.4 & 0.5 \\
0.3 & 0.4 & 0.3 \\
0.4 & 0.2 & 0.2
\end{bmatrix}
\begin{bmatrix}
0.38 \\
0.33 \\
0.27
\end{bmatrix}
= 
\begin{bmatrix}
0.38 \\
0.33 \\
0.27
\end{bmatrix}
$$

(transition matrix $T$)(eigenvector) = (eigenvalue of 1)(eigenvector)