Eigenvalue eigenvector

Instructor: Dr. Avijit Pal

Linear algebra- II (IC152)

We first recall the following theorem:

Theorem

Let $A = [a_{ij}]$ *be an* $n \times n$ *matrix with eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$, *not necessarily distinct. Then* $\det(A) = \prod_{i=1}^{n} \lambda_i$ *and tr*$(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$.

As a consequence of the above Theorem, we will describe the following corollary

Corollary

Suppose A *is a singular matrix. Then* 0 *is an eigen value of* A.

Theorem

If $\lambda_1, \lambda_2, \ldots, \lambda_k$ *are distinct eigenvalues of a matrix* A *with corresponding eigenvectors* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, *then the set* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ *is linearly independent.*

Outline of the proof

- The proof is by induction on the number m of eigenvalues. The result is obviously true if $m = 1$ as the corresponding eigenvector is non-zero and we know that any set containing exactly one non-zero vector is linearly independent.
- Let the result be true for $m,\ 1 \leq m < k$. We prove the result for $m + 1$. We consider the equation

$$c_1x_1 + c_2x_2 + \cdots + c_{m+1}x_{m+1} = \mathbf{0} \tag{1}$$

  for the unknowns $c_1, c_2, \ldots, c_{m+1}$.
- We have

$$\begin{aligned}
\mathbf{0} &= A\mathbf{0} \\
&= A(c_1x_1 + c_2x_2 + \cdots + c_{m+1}x_{m+1}) \\
&= c_1Ax_1 + c_2Ax_2 + \cdots + c_{m+1}Ax_{m+1} \\
&= c_1\lambda_1x_1 + c_2\lambda_2x_2 + \cdots + c_{m+1}\lambda_{m+1}x_{m+1}.
\end{aligned} \tag{2}$$

- From Equations (1) and (2), we get

$$c_2(\lambda_2 - \lambda_1)\mathbf{x}_2 + c_3(\lambda_3 - \lambda_1) + \cdots + c_{m+1}(\lambda_{m+1} - \lambda_1)\mathbf{x}_{m+1} = \mathbf{0}.$$

- This is an equation in m eigenvectors. So, by the induction hypothesis, we have

$$c_i(\lambda_i - \lambda_1) = 0 \quad \text{for} \quad 2 \leq i \leq m+1.$$

- But the eigenvalues are distinct implies $\lambda_i - \lambda_1 \neq 0$ for $2 \leq i \leq m+1$. We therefore get $c_i = 0$ for $2 \leq i \leq m+1$.
- Also, $\mathbf{x}_1 \neq \mathbf{0}$ and therefore (1) gives $c_1 = 0$.
- Thus, we have the required result.

We are thus lead to the following important corollary.

Corollary

The eigenvectors corresponding to distinct eigenvalues of an $n \times n$ matrix A are linearly independent.

Motivation

- Let A be a square matrix of order n and let $T_A : \mathbb{F}^n \longrightarrow \mathbb{F}^n$ be the corresponding linear transformation.
- We ask the question “does there exist a basis $\mathcal{B}$” of $\mathbb{F}^n$ such that $T_A[\mathcal{B}, \mathcal{B}]$, the matrix of the linear transformation T_A, is in the simplest possible form."
- We know that, the simplest form for a matrix is the identity matrix and the diagonal matrix.
- We show that for a certain class of matrices A, we can find a basis $\mathcal{B}$ such that $T_A[\mathcal{B}, \mathcal{B}]$ is a diagonal matrix, consisting of the eigenvalues of A.
- This is equivalent to saying that A is similar to a diagonal matrix. To show the above, we need the following definition.

Definition (Matrix Diagonalisation)

A matrix A is said to be diagonalisable if there exists a non-singular matrix P such that $P^{-1}AP$ is a diagonal matrix.

Remark

Let A be an $n \times n$ diagonalisable matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. By definition, A is similar to a diagonal matrix D. Observe that $D =$ diag$(\lambda_1, \lambda_2, \ldots, \lambda_n)$ as similar matrices have the same set of eigenvalues and the eigenvalues of a diagonal matrix are its diagonal entries.

Example 1

- Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.
- Consider $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Note that $\det P = -2 \neq 0$. Hence, P is invertible.
- Observe that $P^{-1} = \frac{1}{2} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$.
- $AP = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = PD.$
- Since P is invertible, we get $P^{-1}AP = D$.

Example 2

- Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.
- We will show that there does not exist an invertible matrix P such that $P^{-1}AP = D$.
- We will show it by contradiction. Suppose there exist an invertible matrix P such that $P^{-1}AP = D$, that is,

$$P^{-1}AP = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} = D$$

  which is equivalent to say that $A = P \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} P^{-1}$.
- Also,

$$A^2 = P \begin{bmatrix} d_1^2 & 0 \\ 0 & d_2^2 \end{bmatrix} P^{-1}. \tag{3}$$

Example 2 cont.

- Note that

$$A^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{4}$$

- From (3) and (4), we have

$$P^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} P = \begin{pmatrix} d_1^2 & 0 \\ 0 & d_2^2 \end{pmatrix},$$

  which implies $d_1 = d_2 = 0$.

- This shows that $P^{-1}AP = \mathbf{0_{2\times 2}}$, which gives $A = P\,\mathbf{0_{2\times 2}}P^{-1} = \mathbf{0_{2\times 2}}$, which is a contradiction, because $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \mathbf{0_{2\times 2}}$.
- This shows that A is not a diagonalizable matrix.

Consequences of the rank-nullity theorem

The following are some of the consequences of the rank-nullity theorem. The proof is left as an exercise for the reader.

Theorem

The following are equivalent for an $m \times n$ real matrix A.

- *Rank $(A) = k$.*
- *There exist exactly k rows of A that are linearly independent.*
- *There exist exactly k columns of A that are linearly independent.*
- *There is a $k \times k$ submatrix of A with non-zero determinant and every $(k+1) \times (k+1)$ submatrix of A has zero determinant.*
- *The dimension of the range space of A is k.*
- *There is a subset of $\mathbb{R}^m$ consisting of exactly k linearly independent vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$ such that the system $A\mathbf{x} = \mathbf{b}_i$ for $1 \leq i \leq k$ is consistent.*
- *The dimension of the null space of $A = n - k$.*

Definition

A polynomial $f(t)$ in $P(\mathbb{F})$ splits over $\mathbb{F}$ if there are scalars $c, a_1, \cdots, a_n$(not necessarily distinct) in $\mathbb{F}$ such that

$$f(t) = c(t - a_1) \cdots (t - a_n).$$

- For example, $t^2 - 1 = (t + 1)(t - 1)$ splits over $\mathbb{R}$.
- But $(t^2 + 1)(t - 2)$ does not split over $\mathbb{R}$.
- However, $(t^2 + 1)(t - 2)$ splits over $\mathbb{C}$.
- If $f(t)$ is the characteristic polynomial of any matrix over a field $\mathbb{F}$, then the statement that $f(t)$ splits is understood to mean that it splits over $\mathbb{F}$.

Algebraic multiplicity

Theorem

The characteristic polynomial of any diagonalizable matrix splits.

- From the above theorem, it is clear that if A is a $n \times n$ diagonalizable matrix that fails to have distinct eigenvalues, the characteristic polynomial of A must have repeated zeros.

Definition

Let λ be an eigenvalue of a matrix A with characteristic polynomial $f(t)$. The algebraic multiplicity of λ is the largest positive integer k for which $(t-\lambda)^k$ is a factor of $f(t)$.

By using previous theorem, we will prove the following theorem. The following theorem states the necessary and sufficient condition for diagonalizability of a matrix.

Theorem

let A be an $n \times n$ matrix. Then A is diagonalisable if and only if A has n linearly independent eigenvectors.

Outline of the proof

- Let A be diagonalisable. Then there exist matrices P and D such that

$$P^{-1}AP = D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

  Or equivalently, $AP = PD$.

- Let $P = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$. Then $AP = PD$ implies that

$$A\mathbf{u}_i = d_i\mathbf{u}_i \quad \text{for} \quad 1 \leq i \leq n.$$

- Since $\mathbf{u}_i$'s are the columns of a non-singular matrix P, they are non-zero and so for $1 \leq i \leq n$, we get the eigenpairs $(d_i, \mathbf{u}_i)$ of A.
- Since, $\mathbf{u}_i$'s are columns of the non-singular matrix P, using Theorem (6), we get $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are linearly independent.
- Thus we have shown that if A is diagonalisable then A has n linearly independent eigenvectors.

Outline of the proof cont. Converse part

- Conversely, suppose A has n linearly independent eigenvectors $\mathbf{u}_i,\ 1 \leq i \leq n$ with eigenvalues λ_i. Then $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$.
- Let $P = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$. Since $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are linearly independent, by Theorem (6) , P is non-singular.
- Also,

$$\begin{aligned} AP &= [A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n] \\ &= [\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \cdots, \lambda_n\mathbf{u}_n] \\ &= [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\ &= PD. \end{aligned}$$

- Therefore the matrix A is diagonalisable.

As a consequence of the above theorem, we will prove the following lemma.

Corollary

let A be an $n \times n$ matrix. Suppose that the eigenvalues of A are distinct. Then A is diagonalisable.

- As A is an $n \times n$ matrix, it has n eigenvalues. Since all the eigenvalues of A are distinct, by Corollary (4) , the n eigenvectors are linearly independent.
- Hence, by Theorem (10), A is diagonalisable.

Corollary

Let A be an $n \times n$ matrix with $\lambda_1, \lambda_2, \ldots, \lambda_k$ as its distinct eigenvalues and $p(\lambda)$ as its characteristic polynomial. Suppose that for each $i,\ 1 \leq i \leq k,\ (x - \lambda_i)^{m_i}$ divides $p(\lambda)$ but $(x - \lambda_i)^{m_i+1}$ does not divides $p(\lambda)$ for some positive integers m_i . Then

A *is diagonalisable if and only if* $\dim\big(\ker(A - \lambda_i I)\big) = m_i$ *for each* $i,\ 1 \leq i \leq k.$

Or equivalently

A *is diagonalisable if and only if rank*$(A - \lambda_i I) = n - m_i$ *for each* $i,\ 1 \leq i \leq k.$

Outline of the proof

- As A is diagonalisable, by Theorem (10), A has n linearly independent eigenvalues.
- Also, $\sum_{i=1}^{k} m_i = n$ as $\deg(p(\lambda)) = n$.
- Hence, for each eigenvalue $\lambda_i,\ 1 \leq i \leq k$, A has exactly m_i linearly independent eigenvectors.
- Thus, for each $i,\ 1 \leq i \leq k$, the homogeneous linear system $(A - \lambda_i I)\mathbf{x} = \mathbf{0}$ has exactly m_i linearly independent vectors in its solution set.
- Therefore, $\dim\big(\ker(A - \lambda_i I)\big) \geq m_i$. Indeed $\dim\big(\ker(A - \lambda_i I)\big) = m_i$ for $1 \leq i \leq k$ follows from a simple counting argument.

- Now suppose that for each $i, \ 1 \leq i \leq k, \ \dim\big(\ker(A - \lambda_i I)\big) = m_i$.
- Then for each $i, \ 1 \leq i \leq k$, we can choose m_i linearly independent eigenvectors.
- Also by Corollary (4), the eigenvectors corresponding to distinct eigenvalues are linearly independent.
- Hence A has $n = \sum_{i=1}^{k} m_i$ linearly independent eigenvectors.
- Hence by Theorem (10), A is diagonalisable.

Test for diagonalization

Let A be $n \times n$ matrix. Then A is diagonalisable if and only if both of the following conditions hold:

- The characteristic polynomial of A splits.
- For each eigenvalue λ of A, the multiplicity of λ equals $n - \operatorname{rank}(A - \lambda I)$.

Example 1

- Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & -1 & 1 \end{bmatrix}$.
- Then $\det(A - \lambda I) = (2-\lambda)^2(1-\lambda)$. Hence, A has eigenvalues $1, 2, 2$.
- $A - 2I = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}$ has rank 2, we see that $3 - \operatorname{rank}(A - \lambda I) = 1$, which is not multiplicity of 2.
- It is easily seen that $\left(1, (1, 0, -1)^t\right)$ and $(\left(2, (1, 1, -1)^t\right)$ are the only eigenpairs.
- That is, the matrix A has exactly one eigenvector corresponding to the repeated eigenvalue 2.
- Hence, by Theorem (10), the matrix A is not diagonalisable.

Example 2

- Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. Then $\det(A - \lambda I) = (4-\lambda)(1-\lambda)^2$. Hence, A has eigenvalues $1, 1, 4$.
- It can be easily verified that $(1,-1,0)^t$ and $(1,0,-1)^t$ correspond to the eigenvalue 1 and $(1,1,1)^t$ corresponds to the eigenvalue 4.
- Note that the set $\{(1,-1,0)^t, (1,0,-1)^t\}$ consisting of eigenvectors corresponding to the eigenvalue 1 are not orthogonal.
- This set can be replaced by the orthogonal set $\{(1,0,-1)^t, (1,-2,1)^t\}$ which still consists of eigenvectors corresponding to the eigenvalue 1 as $(1,-2,1) = 2(1,-1,0) - (1,0,-1)$.
- Also, the set $\{(1,1,1), (1,0,-1), (1,-2,1)\}$ forms a basis of $\mathbb{R}^3$. So, by Theorem (10), the matrix A is diagonalisable.
- Also, if $U = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}$ is the corresponding unitary matrix then $U^*AU =$ diag$(4,1,1)$.

- Observe that the matrix A is a symmetric matrix. In this case, the eigenvectors are mutually orthogonal.
- In general, for any $n \times n$ real symmetric matrix A, there always exist n eigenvectors and they are mutually orthogonal. This result will be proved later.

Thank You