Linear Algebra

Sartaj Ul Hasan

विद्याधनं सर्वधन प्रधानम्

Department of Mathematics
Indian Institute of Technology Jammu
Jammu, India - 181221

Email: sartaj.hasan@iitjammu.ac.in

Lecture 36

(Nov 01, 2019)

Eigenvectors and Eigenvalues

Eigenvectors and Eigenvalues

- **Definition:** An **eigenvector** of an $n \times n$ matrix $A \in \mathbb{R}^{n \times n}$ is a **non-zero** vector X such that $AX = \lambda X$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a non-trivial solution of $AX = \lambda X$; such a vector X is called an **eigenvector corresponding** to λ.
- Eigenvalues are sometimes also called *characteristic values* or *latent roots*. Eigenvectors are sometimes also called *characteristic vectors*.
- **Remark 1:** The zero vector is **not considered as an eigenvector** since $A0 = \lambda 0$ for all matrices A and all scalars λ.
- **Remark 2:** However, 0 is allowed to be an eigenvalue for a matrix A.

  In that case, the equation $AX = 0X$ has a non-trivial solution. In other words, the equation $AX = 0$ has a non-trivial solution. But $AX = 0$ has a non-trivial solution if and only if A is not invertible. Therefore, an $n \times n$ matrix A is invertible if and only if 0 is **not** an eigenvalue of A.
- **Thus, we have obtained another condition to add to TIMT !!**

Eigenvectors and Eigenvalues (Cont'd)

- **Remark 3:** An eigenvector is not unique, since all scalar multiples of an eigenvector are also eigenvectors. Actually, the set of all eigenvectors corresponding to a particular eigenvalue together with the zero vector forms a subspace of $V = \mathbb{R}^n$ for some n. More formally:
- Put

$$\begin{aligned} E_\lambda &= \{v \in V \ : \ v \ \text{ is an eigenvector for } \ \lambda\} \cup \{0\} \\ &= \{v \ : \ Av = \lambda v\}. \end{aligned}$$

  Then E_λ is a subspace of $V = \mathbb{R}^n$, called the **eigenspace** of A corresponding to λ.
- This can be proved using the subspace test, but follows easily from the fact that the eigenspace corresponding to λ is nothing but the null space of the matrix $(A - \lambda I)$.

Example for eigenvalues and eigenvectors

Let

$$A = \begin{bmatrix} 4 & 2 & -1 \\ -3 & -1 & 1 \\ 6 & 4 & -1 \end{bmatrix}$$

Let $v = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$. Then $Av = \begin{bmatrix} 20 \\ -11 \\ 38 \end{bmatrix}$. So $Av \neq cv$ for any scalar c, and hence v is not an eigenvector.

On the other hand, let $v_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$, $v_2 = \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix}$, $v_3 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$.

Then:

$$Av_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = 1.v_1, \text{ and } Av_2 = \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix} = 1.v_2.$$

Hence, we see that v_1 and v_2 are eigenvectors of A corresponding to the eigenvalue $\lambda_1 = 1$.

Similarly, $Av_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0.v_3$. Therefore, v_3 is an eigenvector corresponding to the eigenvalue $\lambda_2 = 0$ (0 is allowed to be an eigenvalue).

Put $v_4 = v_1 + v_2 = \begin{bmatrix} 3 \\ 1 \\ 11 \end{bmatrix}$. Then:

$$Av_4 = \begin{bmatrix} 4 & 2 & -1 \\ -3 & -1 & 1 \\ 6 & 4 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 11 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 11 \end{bmatrix} = 1.v_4,$$

and hence v_4 is again an eigenvector for λ_1

So if we manage to find one eigenvector for an eigenvalue, we can find more by taking sums and scalar multiples.

Example (Cont'd)

Question: How to check whether a given λ is indeed an eigenvalue for a given matrix A?

Ans: The equation $AX = \lambda X$ or $(A - \lambda I)X = 0$ should have a non-trivial solution X.

Example:

$$A = \begin{bmatrix} 4 & 2 & -1 \\ -3 & -1 & 1 \\ 6 & 4 & -1 \end{bmatrix}$$

We have already seen that 1 and 0 are eigenvalues of A.

How about $\lambda = 3$?

Let us try:

$$A - \lambda I = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -4 & 1 \\ 6 & 4 & -4 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 6R_1]{R_2 \to R_2 + 3R_1} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 2 & -2 \\ 0 & -8 & 2 \end{bmatrix}$$

Example (Cont'd)

$$\xrightarrow[R_3 \to R_3 + 8R_1]{R_2 \to 1/2R_2} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & -14 \end{bmatrix} \xrightarrow{\text{we omit remaining steps,}}$$

but clearly $A - \lambda I$ is row-equivalent to I_3, hence $(A - \lambda I)X = 0$ has only trivial solution. Therefore, 3 is not an eigenvalue.

Fundamental Result about Eigenvectors and Eigenvalues

- **Proposition 40:** If $v_1, v_2, \ldots, v_p$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$, of the matrix A, then the set $\{v_1, v_2, \ldots, v_p\}$ is linearly independent.

  Proof: Suppose BWOC that $v_1, v_2, \ldots, v_p$ are linearly dependent. Let m be the smallest number such that $v_1, v_2, \ldots, v_m$ are linearly independent <u>and</u> v_{m+1} is a linear combination of the preceding vectors (the value of m can be at most $p-1$). Then:

$$c_1 v_1 + c_2 v_2 + \cdots + c_m v_m = v_{m+1} \tag{1}$$

  Apply A on both sides of (1), we get:
  $c_1 A v_1 + c_2 A v_2 + \cdots + c_m A v_m = A v_{m+1}$,
  and using the fact that v_i's are eigenvectors, we get:

$$c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + c_m \lambda_m v_m = \lambda_{m+1} v_{m+1}. \tag{2}$$

- **Proof of Proposition 40 (Cont'd):** Multiplying (1) by λ_{m+1} and subtracting from (2), we get:

$$c_1(\lambda_1 - \lambda_{m+1})v_1 + c_2(\lambda_2 - \lambda_{m+1})v_2 + \cdots + c_m(\lambda_m - \lambda_{m+1})v_m = 0 \quad (3)$$

  However, $v_1, v_2, \ldots, v_m$ are linearly independent, so all of the coefficients in (3) have to be zero, i.e., $c_1(\lambda_1 - \lambda_{m+1}) = 0$ implies $c_1 = 0$, since the λ's are given to be distinct. Similarly, $c_2 = c_3 = \cdots = c_m = 0$. But then from equation (1) we get that $v_{m+1} = 0$, which is a contradiction, since all the v_i's are eigenvectors. Hence our initial hypothesis must be wrong and so, $v_1, v_2, \ldots, v_p$ are linearly independent, and we have the result.
- **Corollary 40.1:** An $n \times n$ matrix A can have at most n distinct eigenvalues.

How to Determine Eigenvalues and Eigenvectors

- **Remark:** It is easy to verify whether a particular vector is an eigenvector of a given matrix A or not. Similarly, given some number, we can verify whether it is an eigenvalue or not.
- However, in order to systematically find eigenvalues, we use the following result:
- **Proposition 41:** A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if it satisfies the characteristic equation $\det(A - \lambda I) = 0$.
- **Note:** $\det(A - \lambda I)$ is a polynomial of degree n called the **characteristic polynomial** of A. It has at most n roots, counting multiplicities. Hence an $n \times n$ matrix can have at most n eigenvalues (counting multiplicities). It is possible for a matrix with real entries to have no real eigenvalues.
- **Note:** If complex roots are allowed, an $n \times n$ matrix has exactly n eigenvalues (counting multiplicities). Therefore, we must clearly specify which field is being considered when we talk about the eigenvalues of a matrix. For the time being, however, we will only allow real eigenvalues.

Proof of Proposition 41

$[\implies]$

Suppose λ is an eigenvalue of A. Then there is a non-zero vector v such that $Av = \lambda v$ (v is an eigenvector), i.e., $(A - \lambda I)v = 0$. Therefore, the homogeneous system $(A - \lambda I)X = 0$ has a non-zero solution, which implies that the matrix $(A - \lambda I)$ is not invertible and as a consequence, $\det(A - \lambda I) = 0$.

$[\impliedby]$

Conversely, suppose that λ is a root of the characteristic equation, i.e., $\det(A - \lambda I) = 0$. Therefore, the matrix $(A - \lambda I)$ is not invertible. Thus, by TIMT, the homogeneous system $(A - \lambda I)X = 0$ has a non-zero solution, say v, i.e., $(A - \lambda I)v = 0$, which implies that $Av = \lambda v$, and hence, λ is an eigenvalue of A.

Eigenvalues of Similar Matrices

- Recall that an $n \times n$ matrix B is said to be similar to an $n \times n$ matrix A if there exists an invertible matrix P such that $B = PAP^{-1}$ (or $A = P^{-1}BP$). Similarity of matrices is an equivalence relation on the set of $n \times n$ matrices.
- **Remark:** Using the multiplicative property of determinants, it is easy to see that similar matrices have the same determinant. Using essentially the same idea, we can derive the following result:
- **Proposition 42:** If the $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial, and hence the same eigenvalues with the same multiplicities.

Proof of Proposition 42:

Suppose B is similar to A, i.e., $B = PAP^{-1}$ for some invertible matrix P.

$$
\begin{aligned}
\therefore \quad \text{char. poly. of } \; B &= \det(B - \lambda I) \\
&= \det(PAP^{-1} - \lambda I) \\
&= \det(PAP^{-1} - P(\lambda I)P^{-1}) \\
&= \det(P(A - \lambda I)P^{-1}) \\
&= \det P \det(A - \lambda I) \det(P^{-1}) \\
&= \det P \det(A - \lambda I)(\det(P))^{-1} \\
&= \det(A - \lambda I).
\end{aligned}
$$

Remark: Converse of Proposition 42 is not true! We can find matrices A and B such that char. poly of $A =$ char. poly. of B, but B is not similar to A, e.g, the matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ have same char poly, but they are not similar.

Diagonalization of Matrices

Diagonalization of Matrices

- If A is a diagonal matrix, then its diagonal elements are its eigenvalues, and the standard basis vectors are its eigenvectors. This is the motivation for the following:
- **Definition:** An $n \times n$ matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix D. In other words, if there is an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.
- **Remark 1:** If A is diagonalizable, then its powers are easy to compute.
- **Remark 2:** If A is diagonalizable, then its eigenvalues can be found by inspection of D. However, in practice, we have to do things the other way round. First, we find the eigenvalues from the characteristic equation, then we find P with the help of the corresponding eigenvectors, then we get the diagonal matrix D.

Diagonalization of Matrices (Cont'd)

- **Theorem 4 (Diagonalization Theorem (DT)):**
  a. An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
  b. In this case, $A = PDP^{-1}$, where the columns of P are n linearly independent eigenvectors of A, and the diagonal entries of D are eigenvalues corresponding to these eigenvectors.

Diagonalization of Matrices (Cont'd)

- **Theorem 4 (Diagonalization Theorem (DT)):**
  - (a) An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
  - (b) In this case, $A = PDP^{-1}$, where the columns of P are n linearly independent eigenvectors of A, and the diagonal entries of D are eigenvalues corresponding to these eigenvectors.

Proof of part (a): $[\implies]$
Suppose A is diagonalizable. Then $A = PDP^{-1}$ for some diagonal matrix D and some invertible matrix P, i.e.,

$$AP = PD \tag{4}$$

Let $P = [v_1, v_2, \ldots, v_n]$ and let $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where the λ's need not be distinct. Therefore (4) becomes

$$A[v_1, v_2, \ldots, v_n] = P \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

- **Proof of Part (a) (Cont'd):**

$$[Av_1, Av_2, \ldots, Av_n] = [v_1, v_2, \ldots, v_n] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$
$$= [\lambda_1 v_1, \lambda_2 v_2, \ldots, \lambda_n v_n]$$

Equating columns, we get:

$$Av_i = \lambda_i v_i, i = 1, 2, \ldots, n \quad (5)$$

Now, the vectors $v_1, \ldots, v_n$ being columns of an invertible matrix are linearly independent and (5) shows that they are eigenvectors.

$[\Longleftarrow]$
Conversely, suppose A has n linearly independent eigenvectors so that $Av_i = \lambda_i v_i, i = 1, 2, \ldots, n$. Form the matrix P with v_i's as columns.

- **Proof of Part (a) (Cont'd):**
  Then:

$$\begin{aligned} AP &= A[v_1, v_2, \dots, v_n] \\ &= [Av_1, Av_2, \dots, Av_n] \\ &= [\lambda_1 v_1, \lambda_2 v_2, \dots, \lambda_n v_n] \\ &= PD, \text{ where } D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n). \end{aligned}$$

  But P is invertible, so $A = PDP^{-1}$, and A is diagonalizable.
- **Proof of Part (b):** Part (b) has been proved en route to proving part (a).
- **Remark:** Another way to express the above theorem is that an $n \times n$ matrix A is diagonalizable if and only if it has enough (linearly independent) eigenvectors to form a basis of $\mathbb{R}^n$. Such a basis is called an **eigenvector basis**.