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 37

(Nov 05, 2019)

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)):**
  1. An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
  2. 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)):**
  1. (a) An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
  2. (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.,

$$AP = PD \tag{1}$$

Let $P = [v_1, v_2, \dots, v_n]$ and let $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ where the λ's need not be distinct. Therefore (1) becomes

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

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

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

Equating columns, we get:

$$Av_i = \lambda_i v_i, i = 1, 2, \dots, n \tag{2}$$

Now, the vectors $v_1, \dots, v_n$ being columns of an invertible matrix are linearly independent and (2) shows that they are eigenvectors.

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

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

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

  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**.

Diagonalization of Matrices (Cont'd)

- In practice, we can distinguish three cases:
- **Case 1:** An $n \times n$ matrix A has n distinct (real) eigenvalues. Then we get the following result:
- **Proposition 43:** An $n \times n$ matrix A with n distinct eigenvalues is diagonalizable.

  Proof: By an earlier result (Proposition 40), eigenvectors corresponding to distinct eigenvalues are linearly independent. Therefore in this case A has n linearly independent eigenvectors. Hence by the DT, A is diagonalizable.

Example for Case 1

$$A = \begin{bmatrix} 42 & -33 \\ 22 & -13 \end{bmatrix}$$

Then:

$$\begin{aligned}
\det(A - \lambda I) &= \det \begin{bmatrix} 42 & -33 \\ 22 & -13 \end{bmatrix} \\
&= (42 - \lambda)(-13 - \lambda) + 22(33) \\
&= 180 - 29\lambda + \lambda^2 \\
&= (20 - \lambda)(9 - \lambda)
\end{aligned}$$

Hence there are 2 distinct eigenvalues $\lambda_1 = 20, \lambda_2 = 9$.

(i) For $\lambda_1 = 20$, $A - \lambda_1 I = \begin{bmatrix} 22 & -33 \\ 22 & -33 \end{bmatrix} \to \begin{bmatrix} 22 & -33 \\ 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -3/2 \\ 0 & 0 \end{bmatrix}$ or

$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 3/2 \\ 1 \end{bmatrix}$.

Example for Case 1 (Cont'd)

Let us take $v_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ as an eigenvector.

[**Check:** $Av_1 = \begin{bmatrix} 42 & -33 \\ 22 & -13 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 60 \\ 40 \end{bmatrix} = 20v_1$ as required.]

(ii) For $\lambda_2 = 9$, $A - \lambda_2 I = \begin{bmatrix} 33 & -33 \\ 22 & -22 \end{bmatrix} \to \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$. Therefore,

$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, so we take $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as the eigenvector.

[**Check:** $Av_2 = \begin{bmatrix} 42 & -33 \\ 22 & -13 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 9 \end{bmatrix} = 9v_2$ as required.]

Note: We should get $A = PDP^{-1}$ where $D = \text{diag}(20, 9)$ and $P = [v_1, v_2]$. Instead of finding P^{-1}, it is easier to check $AP = PD$.

$$AP = \begin{bmatrix} 42 & -33 \\ 22 & -13 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 60 & 9 \\ 40 & 9 \end{bmatrix}$$

$$PD = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 20 & 0 \\ 0 & 9 \end{bmatrix} = \begin{bmatrix} 60 & 9 \\ 40 & 9 \end{bmatrix}$$

Diagonalization of Matrices (Cont'd)

- **Two Preliminary Definitions:** Given an eigenvalue λ_1 for a matrix A, we define:
  - The **algebraic multiplicity** of λ_1 is the power of the factor $(\lambda - \lambda_1)$ in the characteristic polynomial of A.
  - The **geometric multiplicity** of λ_1 is the dimension of the eigenspace corresponding to λ_1.
- **Remark:** The first definition is the one we have used for multiplicity up to now, as it applies to polynomials in general (not only the characteristic polynomial). The second definition applies specifically to the characteristic polynomial, since its roots are eigenvalues, which have corresponding eigenspaces.

Diagonalization of Matrices (Cont'd)

- **Case 2:** An $n \times n$ matrix A has $p < n$ distinct eigenvalues, but counting the (algebraic) multiplicities, there are n real eigenvalues (not distinct). We now come to the weaker result for this case:
- **Proposition 44:** Let A be an $n \times n$ matrix with n (real) eigenvalues (counting algebraic multiplicities) of which only $\lambda_1, \lambda_2, \ldots, \lambda_p$ are distinct ($p < n$). Then the following hold:
  - (a) For $1 \leq k \leq p$, the geometric multiplicity is less than or equal to the algebraic multiplicity of λ_k.
  - (b) A is diagonalizable if and only if the sum of the dimensions of the distinct eigenspaces is n, and this happens if and only if the geometric multiplicity for each λ_k equals its algebraic multiplicity.
  - (c) If A is diagonalizable, and B_k is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in $B_1, B_2, \ldots, B_p$ forms an eigenvector basis for $\mathbb{R}^n$.

Examples of Case 2

Example 1:

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

Char poly of $A = \det(A - \lambda I) = (-\lambda)(1-\lambda)^2$
Therefore, the eigenvalues are:

$$\lambda_1 = 1 \quad \text{(alg. multiplicity 2)}$$

$$\lambda_2 = 0 \quad \text{(alg. multiplicity 1)}$$

(i) Taking $\lambda_1 = 1$: $A - \lambda_1 I = \begin{bmatrix} 3 & 2 & -1 \\ -3 & -2 & 1 \\ 6 & -4 & -2 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 2R_1]{R_2 \to R_2 + R_1} \begin{bmatrix} 3 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Divide the first row by 3, we would get the following solution.

Example 1 of Case 2 (Cont'd)

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2/3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1/3 \\ 0 \\ 1 \end{bmatrix}$$

Putting $x_2 = -3, x_3 = 0$, we get $v_1 = \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$ as an eigenvector.

Putting $x_2 = 0, x_3 = 3$, we get $v_2 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ as an eigenvector.

(ii) Taking $\lambda_2 = 0$:

$$A - \lambda_1 I = \begin{bmatrix} 4 & 2 & -1 \\ -3 & -1 & 1 \\ 6 & 4 & -1 \end{bmatrix} \xrightarrow{R_1 \to R_1 + R_2} \begin{bmatrix} 1 & 1 & 0 \\ -3 & -1 & 1 \\ 6 & 4 & -1 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 6R_1]{R_2 \to R_2 + 3R_1}$$

Example 1 of Case 2 (Cont'd)

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{bmatrix} \xrightarrow{R_3 \to R_3 + R_2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \to (1/2) R_2}$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \to R_1 - R_2} \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$. Putting $x_3 = 2$, we get $v = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ as an eigen value corresponding to λ_2. In this case, A has turned out to be diagonalizable, because geometric multiplicity of $\lambda_1 = 2 =$ algebraic multiplicity of λ_1.

Note: If algebraic multiplicity of an eigenvalue equals 1, then geometric multiplicity has to be 1. So the problem of diagonalization would arise only if algebraic multiplicity of an eigenvalue is > 1.

Example 2 of Case 2

Example 2:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 0 \end{bmatrix}$$

Char poly of $A = \det(A - \lambda I) = (-\lambda)(1-\lambda)^2$
So we get the same situation as in Example 1:

$$\lambda_1 = 1 \quad \text{(alg. multiplicity 2)}$$

$$\lambda_2 = 0 \quad \text{(alg. multiplicity 1)}$$

(i) Taking $\lambda_1 = 1$**:** $A - \lambda_1 I = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 2 & -1 \end{bmatrix} \xrightarrow{\text{interchange}} \begin{bmatrix} 1 & 0 & 0 \\ 3 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

Example 2 of Case 2 (Cont'd)

$$\xrightarrow{R_2 \to R_2 - 3R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \to (1/2)R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 1/2 \\ 1 \end{bmatrix}$; putting $x_3 = 2$, we get $v_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$.

v_1 is certainly an eigenvector, since

$$Av_1 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = 1.v_1$$

However, geometric multiplicity of $\lambda_1 (= 1) <$ algebraic multiplicity $(= 2)$. Therefore, A is NOT diagonalizable.

Note: If geometric multiplicity $<$ algebraic multiplicity for any one eigenvalue , the matrix is **not** diagonalizable.

Analysis of Case 2

We have $\lambda_1, \ldots, \lambda_p$ as distinct real eigenvalues, where $p < n$, but algebraic multiplicities add up to n.

$$
\begin{array}{ccc}
\text{geom mult } \lambda_1 & \leq & \text{alg mult } \lambda_1 \\
\text{geom mult } \lambda_2 & \leq & \text{alg mult } \lambda_2 \\
\vdots & & \vdots \\
\text{geom mult } \lambda_p & \leq & \text{alg mult } \lambda_p \\
------- & & ------- \\
n & & n
\end{array}
$$

If A is diagonalizable, then LHS must also add to n (by DT).

So if part (a) of Proposition 44 is proved, then (b) has to follow – i.e., all geom multiplicities must equal the corresponding alg multiplicities. (c) follows without much trouble. But (a) is advanced and beyond our scope.

Diagonalization of Matrices (Cont'd)

- **Case 3:** An $n \times n$ matrix A has $p < n$ distinct eigenvalues, but even after counting the algebraic multiplicities, there are $< n$ real eigenvalues (p could even be 0). Then A is not diagonalizable over the real field. If we want to diagonalize, we have to admit complex eigenvalues and eigenvectors.
- **Remark:** Even if we admit complex eigenvalues and eigenvectors, a real matrix does not have to be diagonalizable. The case is quite complicated, and we will not go into the details.