

MA 110

Linear Algebra and Differential Equations

Lecture 12

Prof. Sudhir R. Ghorpade
Department of Mathematics
IIT Bombay

<http://www.math.iitb.ac.in/~srg/>

Spring 2025

Similarity and Eigenvalues

Recall that $\lambda \in \mathbb{K}$ is an **eigenvalue** of $\mathbf{A} \in \mathbb{K}^{n \times n}$ if $\mathbf{Ax} = \lambda \mathbf{x}$ for some $\mathbf{x} \in \mathbb{K}^{n \times 1}$ with $\mathbf{x} \neq \mathbf{0}$. It turns out that similar matrices have the same eigenvalues. In fact, more is true.

Proposition

Let $\mathbf{A}, \mathbf{A}' \in \mathbb{K}^{n \times n}$ be similar. Then $p_{\mathbf{A}}(t) = p_{\mathbf{A}'}(t)$. In particular, $\lambda \in \mathbb{K}$ is an eigenvalue of $\mathbf{A}$ if and only if λ is an eigenvalue of $\mathbf{A}'$. Consequently, the algebraic multiplicity of λ as an eigenvalue of $\mathbf{A}$ is equal to the algebraic multiplicity of λ as an eigenvalue of $\mathbf{A}'$. Furthermore, the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}$ is equal to the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}'$.

Proof: Since $\mathbf{A} \sim \mathbf{A}'$, there is an invertible $\mathbf{P} \in \mathbb{K}^{n \times n}$ such that $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Writing $\mathbf{I} = \mathbf{P}^{-1}\mathbf{P}$, we see that

$$p_{\mathbf{A}'}(t) = \det(\mathbf{A}' - t\mathbf{I}) = \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - t\mathbf{I}) = \det(\mathbf{A} - t\mathbf{I}) = p_{\mathbf{A}}(t).$$

Proof Contd. It remains to prove the assertion about geometric multiplicity. Let λ be an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ corresponding to λ . Then $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Since $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, we see that $\mathbf{x}' := \mathbf{P}^{-1}\mathbf{x}$ satisfies $\mathbf{A}'\mathbf{x}' = \lambda\mathbf{x}'$. Also $\mathbf{x}' \neq \mathbf{0}$ since $\mathbf{P}$ is invertible. Thus $\mathbf{x}'$ is an eigenvector of $\mathbf{A}'$ corresponding to λ . Also, it is easy to check that $\{\mathbf{x}_1, \dots, \mathbf{x}_g\}$ is a basis for $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})$ if and only if $\{\mathbf{P}^{-1}\mathbf{x}_1, \dots, \mathbf{P}^{-1}\mathbf{x}_g\}$ is a basis for $\mathcal{N}(\mathbf{A}' - \lambda\mathbf{I})$. Hence the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}$ is equal to the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}'$. $\square$

Examples:

$$(i) \mathbf{A} := \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \Rightarrow p_{\mathbf{A}}(t) = \det \begin{bmatrix} \lambda - t & 1 \\ 0 & \lambda - t \end{bmatrix} = (\lambda - t)^2.$$

Hence λ is the only eigenvalue of $\mathbf{A}$, and its algebraic multiplicity is 2. But its geometric multiplicity is 1 since

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \implies \text{rank}(\mathbf{A} - \lambda\mathbf{I}) = 1 \implies \text{nullity}(\mathbf{A} - \lambda\mathbf{I}) = 1.$$

Note that in the above example, if $\mathbf{A}$ were similar to a diagonal matrix $\mathbf{D}$, then we must have $\mathbf{D} = \text{diag}(\lambda, \lambda)$, since eigenvalues and their algebraic multiplicities of $\mathbf{A}$ and $\mathbf{D}$ have to be the same. But the geometric multiplicity of λ as an eigenvalue of $\mathbf{D}$ is 2. This shows that $\mathbf{A}$ is not diagonalizable.

$$(ii) \mathbf{A} := \begin{bmatrix} 3 & 0 & 0 \\ -2 & 4 & 2 \\ -2 & 1 & 5 \end{bmatrix} \Rightarrow p_{\mathbf{A}}(t) = \det \begin{bmatrix} 3-t & 0 & 0 \\ -2 & 4-t & 2 \\ -2 & 1 & 5-t \end{bmatrix}.$$

Computing the determinant, we find $p_{\mathbf{A}}(t) = (3-t)^2(6-t)$. Hence 3 is an eigenvalue of $\mathbf{A}$ of algebraic multiplicity 2, and 6 is an eigenvalue of $\mathbf{A}$ of algebraic multiplicity 1. Also,

$$\mathbf{A} - 3\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 1 & 2 \\ -2 & 1 & 2 \end{bmatrix} \Rightarrow \text{rank}(\mathbf{A} - 3\mathbf{I}) = 1.$$

So $\text{nullity}(\mathbf{A} - 3\mathbf{I}) = 2$. In fact, $\{ [1 \ 0 \ 1]^T, [1 \ 2 \ 0]^T \}$ is a basis of the eigenspace of $\mathbf{A}$ corresponding to eigenvalue 3, and so its geometric multiplicity is equal to 2.

Similarity and Change of Basis

Similarity of matrices has the following characterisation.

Proposition

Let $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$. Then $\mathbf{A} \sim \mathbf{B}$ if and only if there is an ordered basis E for $\mathbb{K}^{n \times 1}$ such that $\mathbf{B}$ is the matrix of the linear transformation $T_{\mathbf{A}} : \mathbb{K}^{n \times 1} \rightarrow \mathbb{K}^{n \times 1}$ with respect to E .

In fact, $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ if and only if the columns of $\mathbf{P}$ form an ordered basis, say E , for $\mathbb{K}^{n \times 1}$ and $\mathbf{B} = \mathbf{M}_E^E(T_{\mathbf{A}})$.

Proof. Let $\mathbf{B} := [b_{jk}]$. Now $\mathbf{A} \sim \mathbf{B} \iff$ there is an invertible matrix $\mathbf{P}$ such that $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{B}$. This is the case if and only if there is an ordered basis $E := (\mathbf{x}_1, \dots, \mathbf{x}_n)$ for $\mathbb{K}^{n \times 1}$ such that

$$\mathbf{A} \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}.$$

The k th column of LHS is $\mathbf{A}\mathbf{x}_k$ and the k th column of RHS is the linear combination of $\mathbf{x}_1, \dots, \mathbf{x}_n$ with coefficients from the k th column of $\mathbf{B}$. Thus $\mathbf{A}\mathbf{x}_k = b_{1k}\mathbf{x}_1 + \dots + b_{nk}\mathbf{x}_n$ for $k = 1, \dots, n$. This means the k th column of $\mathbf{M}_E^E(T_{\mathbf{A}})$ is the k th column $[b_{1k} \ \dots \ b_{nk}]^T$ of $\mathbf{B}$, $k=1, \dots, n$, that is, $\mathbf{B} = \mathbf{M}_E^E(T_{\mathbf{A}})$. □

The above result says that just as $\mathbf{A}$ is the matrix of the linear transformation $T_{\mathbf{A}}$ defined by $\mathbf{A}$ with respect to the standard ordered basis $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ for $\mathbb{K}^{n \times 1}$, the matrix $\mathbf{B} := \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is the matrix of the same linear transformation $T_{\mathbf{A}}$ with respect to the ordered basis for $\mathbb{K}^{n \times 1}$ consisting of the columns of $\mathbf{P}$.

Now we shall make precise what we mean by ‘a matrix $\mathbf{A}$ behaves like a diagonal matrix’.

Definition

A matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ is called **diagonalizable** (over $\mathbb{K}$) if $\mathbf{A}$ is similar to a diagonal matrix (over $\mathbb{K}$).

Characterization of Diagonalizability

Recall the definition.

Definition

A matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ is called **diagonalizable** (over $\mathbb{K}$) if $\mathbf{A}$ is similar to a diagonal matrix (over $\mathbb{K}$).

We stated the following characterization of diagonalizability.

Proposition

A matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ is diagonalizable if and only if there is a basis for $\mathbb{K}^{n \times 1}$ consisting of eigenvectors of $\mathbf{A}$. In fact,

$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, where $\mathbf{P}, \mathbf{D} \in \mathbb{K}^{n \times n}$ are of the form

$\mathbf{P} = [\mathbf{x}_1 \ \cdots \ \mathbf{x}_n]$ and $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$

$\iff \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is a basis for $\mathbb{K}^{n \times 1}$ and

$\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$ for $k = 1, \dots, n$.

Proof of a characterization of diagonalizability

Proof. The result is a consequence of the earlier characterization of similarity. It can also be seen as follows.

A is diagonalizable

$\iff \exists$ invertible matrix **P** and diagonal matrix **D** in $\mathbb{K}^{n \times n}$ such that **AP** = **PD**

$\iff \exists$ a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ for $\mathbb{K}^{n \times 1}$ and $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ such that $\mathbf{A} \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \text{diag}(\lambda_1, \dots, \lambda_n)$

$\iff \exists$ a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ of $\mathbb{K}^{n \times 1}$ and $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ such that $\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$ for $k = 1, \dots, n$. $\square$

Application: If a matrix **A** is diagonalizable and we find invertible **P** such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$, then any power of **A** can be found easily. This is seen as follows:

$$\mathbf{A}^m = (\mathbf{P}\mathbf{D}\mathbf{P}^{-1}) \cdots (\mathbf{P}\mathbf{D}\mathbf{P}^{-1}) = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1} = \mathbf{P} \text{diag}(\lambda_1^m, \dots, \lambda_n^m)\mathbf{P}^{-1}.$$

Example: Consider $\mathbf{A} := \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$. Then

$$p_{\mathbf{A}}(t) = \det \begin{bmatrix} t-4 & 3 \\ -2 & t+1 \end{bmatrix} = (t-4)(t+1)+6 = (t-2)(t-1).$$

Thus 2 and 1 are the eigenvalues of $\mathbf{A}$ and it is easy to see that $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are corresponding eigenvectors. So $\mathbf{P} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$ satisfies $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(2, 1)$, which can be written as

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}.$$

Hence

$$\begin{aligned} \mathbf{A}^m &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^m & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 2^m 3 & 1 \\ 2^m 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 2^m 3 - 2 & -2^m 3 + 3 \\ 2^m 2 - 2 & -2^m 2 + 3 \end{bmatrix} \quad \text{for } m \in \mathbb{N}. \end{aligned}$$

Eigenvectors corresponding to distinct eigenvalues

Our next result is about the linear independence of eigenvectors corresponding to distinct eigenvalues of a matrix.

Lemma

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of $\mathbf{A}$. Let $\mathbf{x}_1, \dots, \mathbf{x}_k$ belong to the eigenspaces of $\mathbf{A}$ corresponding to $\lambda_1, \dots, \lambda_k$ respectively. Then

$$\mathbf{x}_1 + \dots + \mathbf{x}_k = \mathbf{0} \iff \mathbf{x}_1 = \dots = \mathbf{x}_k = \mathbf{0}.$$

In particular, if $\mathbf{x}_1, \dots, \mathbf{x}_k$ are eigenvectors of $\mathbf{A}$ corresponding to $\lambda_1, \dots, \lambda_k$ respectively, then the set $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ is linearly independent.

Proof. We use induction on the number k of distinct eigenvalues of $\mathbf{A}$. Clearly, the result holds for $k = 1$.

Let $k \geq 2$ and assume that the result holds for $k - 1$.

Suppose $\mathbf{x} := \mathbf{x}_1 + \cdots + \mathbf{x}_{k-1} + \mathbf{x}_k = \mathbf{0}$. Then $\mathbf{A}\mathbf{x} = \mathbf{0}$, that is, $\lambda_1\mathbf{x}_1 + \cdots + \lambda_{k-1}\mathbf{x}_{k-1} + \lambda_k\mathbf{x}_k = \mathbf{0}$. Also, multiplying the first equation by λ_k , we obtain $\lambda_k\mathbf{x}_1 + \cdots + \lambda_k\mathbf{x}_{k-1} + \lambda_k\mathbf{x}_k = \mathbf{0}$. Subtraction gives $(\lambda_1 - \lambda_k)\mathbf{x}_1 + \cdots + (\lambda_{k-1} - \lambda_k)\mathbf{x}_{k-1} = \mathbf{0}$.

By the induction hypothesis,

$(\lambda_1 - \lambda_k)\mathbf{x}_1 = \cdots = (\lambda_{k-1} - \lambda_k)\mathbf{x}_{k-1} = \mathbf{0}$. Since $\lambda_1 \neq \lambda_k, \dots, \lambda_{k-1} \neq \lambda_k$, we obtain $\mathbf{x}_1 = \cdots = \mathbf{x}_{k-1} = \mathbf{0}$, and so $\mathbf{x}_k = \mathbf{0}$ as well.

Now let $\mathbf{x}_1, \dots, \mathbf{x}_k$ be eigenvectors. If $\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}$, then $\alpha_1\mathbf{x}_1 = \cdots = \alpha_k\mathbf{x}_k = \mathbf{0}$. But $\mathbf{x}_1 \neq \mathbf{0}, \dots, \mathbf{x}_k \neq \mathbf{0}$, so that $\alpha_1 = \cdots = \alpha_k = 0$. Thus $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ is linearly independent.

Theorem

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of $\mathbf{A}$. Let g_j be the geometric multiplicity of λ_j for $j = 1, \dots, k$. Then $g_1 + \cdots + g_k \leq n$. Further, $\mathbf{A}$ is diagonalizable if and only if $g_1 + \cdots + g_k = n$.

Proof. Let V_j denote the eigenspace $\mathcal{N}(\mathbf{A} - \lambda_j \mathbf{I})$ of $\mathbb{K}^{n \times 1}$, and let E_j be a basis for V_j consisting of g_j eigenvectors of $\mathbf{A}$ corresponding to λ_j for $j = 1, \dots, k$.

We claim that the set $E := E_1 \cup \dots \cup E_k$ is linearly independent. Let $\mathbf{x}$ be a linear combination of elements of E . Collate the elements of E_j for each $j = 1, \dots, k$ and write $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_k$, where $\mathbf{x}_j \in V_j$ for $j = 1, \dots, k$. Suppose $\mathbf{x} = \mathbf{0}$. Then $\mathbf{x}_j = \mathbf{0}$ for $j = 1, \dots, k$ by the previous lemma. For $j \in \{1, \dots, k\}$, $\mathbf{x}_j$ is a linear combination of elements of the set E_j , and since the set E_j is linearly independent, every coefficient in this linear combination must be 0. Since this holds for each $j = 1, \dots, k$, we see that every coefficient in the linear combination $\mathbf{x}$ of elements of E must be 0. Hence O.K.

The number of elements in the linearly independent set E is $g_1 + \dots + g_k$. Since E is a subset of the n dimensional vector space $\mathbb{K}^{n \times 1}$, it follows that $g_1 + \dots + g_k \leq n$.

Now suppose $g_1 + \cdots + g_k = n$. Then E is a linearly independent subset of $\mathbb{K}^{n \times 1}$ having n elements. Thus E is a basis for $\mathbb{K}^{n \times 1}$ consisting of eigenvectors of $\mathbf{A}$. Hence $\mathbf{A}$ is diagonalizable.

Conversely, suppose $\mathbf{A}$ is diagonalizable. Then there is a basis for $\mathbb{K}^{n \times 1}$ consisting of n eigenvectors of $\mathbf{A}$. For $j = 1, \dots, k$, let h_j elements of this basis belong to V_j . Since these elements form a linearly independent subset of V_j , we see that $h_j \leq g_j$ for $j = 1, \dots, k$. Hence $n = h_1 + \cdots + h_k \leq g_1 + \cdots + g_k \leq n$. This shows that $g_1 + \cdots + g_k = n$. $\square$

Corollary

If $\mathbf{A} \in \mathbb{K}^{n \times n}$ has n distinct eigenvalues, then $\mathbf{A}$ is diagonalizable.

Proof. Clearly, $n = 1 + \cdots + 1 \leq g_1 + \cdots + g_n \leq n$, and so $g_1 + \cdots + g_n = n$. Hence the above theorem applies. $\square$

The case of $\mathbb{K} = \mathbb{C}$

In case $\mathbb{K} = \mathbb{C}$, then by the [Fundamental Theorem of Algebra](#), every polynomial of degree n with coefficients in $\mathbb{C}$ has exactly n roots in $\mathbb{C}$, counting multiplicities. In particular, the characteristic polynomial $p_{\mathbf{A}}(t)$ of any $\mathbf{A} \in \mathbb{C}^{n \times n}$ has exactly n roots in $\mathbb{C}$, counting multiplicities. More precisely, we can factor

$$p_{\mathbf{A}}(t) = (\lambda_1 - t)^{m_1} \cdots (\lambda_k - t)^{m_k},$$

where $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ are distinct and $m_1, \dots, m_k \in \mathbb{N}$ satisfy $m_1 + \cdots + m_k = n$.

As an immediate consequence, we obtain the following result.

Theorem

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$. Then there are distinct eigenvalues $\lambda_1, \dots, \lambda_k$ of $\mathbf{A}$ having algebraic multiplicities $m_1, \dots, m_k$ such that $m_1 + \cdots + m_k = n$.

Proposition

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of $\mathbf{A}$. Let g_j and m_j be the geometric multiplicity and the algebraic multiplicity of λ_j respectively for $j = 1, \dots, k$.

- (i) If $\mathbf{A}$ be diagonalizable, then $g_j = m_j$ for $j = 1, \dots, k$.
- (ii) If $\mathbb{K} = \mathbb{C}$, and $g_j = m_j$ for $j = 1, \dots, k$, then $\mathbf{A}$ is diagonalizable.

Proof. (i) Since $\mathbf{A}$ is diagonalizable, $g_1 + \dots + g_k = n$. Hence

$$0 \leq (m_1 - g_1) + \dots + (m_k - g_k) \leq n - n = 0.$$

But $m_j - g_j \geq 0$, and so $g_j = m_j$ for $j = 1, \dots, k$.

(ii) Since $\mathbb{K} = \mathbb{C}$, $m_1 + \dots + m_k = n$. Also, since $g_j = m_j$ for $j = 1, \dots, k$, we see that $g_1 + \dots + g_k = m_1 + \dots + m_k = n$. Hence $\mathbf{A}$ is diagonalizable. □

Remark

Part (ii) of the above proposition does not hold if $\mathbb{K} = \mathbb{R}$.

For example, let $\mathbf{A} := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$. Then

$p_{\mathbf{A}}(t) = (1 - t)(1 + t^2)$, and the geometric multiplicity as well as the algebraic multiplicity of the only (real) eigenvalue 1 of $\mathbf{A}$ is equal to 1. Thus the 3×3 matrix $\mathbf{A}$ is not diagonalizable (over $\mathbb{R}$) since the sum of the geometric multiplicities of its eigenvalues is less than 3.

On the other hand, if $\mathbb{K} = \mathbb{C}$, then

$p_{\mathbf{A}}(t) = (1 - t)(t - i)(t + i)$, and for each of the eigenvalues $1, i, -i$ of $\mathbf{A}$, the geometric multiplicity as well as the algebraic multiplicity is equal to 1, and so $\mathbf{A}$ is diagonalizable (over $\mathbb{C}$).