MA 110
Linear Algebra and Differential Equations
Lecture 11

Prof. Sudhir R. Ghorpade
Department of Mathematics
IIT Bombay
http://www.math.iitb.ac.in/~srg/

Spring 2025

Recall: The last time we discussed linear transformation $T : V \to W$ where V, W are vector (sub)spaces. If $\dim V = n$ and $\dim W = m$ and if $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $F := (\mathbf{y}_1, \ldots, \mathbf{y}_m)$ are ordered basis of V and W, respectively, then we can associate to T the $m \times n$ matrix $\boxed{\mathbf{M}_F^E(T)}$; it is called the **matrix of T with respect to the ordered bases E and F**. For each $k = 1, \ldots, , n$, the kth column of this matrix consists of the coefficients of $T(\mathbf{x}_k)$ when expressed as a linear combination of $\mathbf{y}_1, \ldots, \mathbf{y}_m$.

We also discussed complex numbers and noted that they could be defined using certain 2×2 matrices with entries in $\mathbb{R}$. The set of all complex numbers is denoted by $\mathbb{C}$ and we regard $\mathbb{R}$ as a subset of $\mathbb{C}$.

Let $z \in \mathbb{C}$. Then $z = x + iy$ for unique $x, y \in \mathbb{R}$. Then x is called the **real part** of z, and it is denoted by $\Re(z)$, while y is called the **imaginary part** of z, and it is denoted by $\Im(z)$.

Let $z \in \mathbb{C}$. Then $z = x + iy$ for unique $x, y \in \mathbb{R}$. The **conjugate** of z is defined to be the complex number $\overline{z} := x - iy$.

We define the **absolute value** of $z = x + iy$ by

$$|z| := \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}.$$

Note that the following **triangle inequality** holds.

$$|z_1 + z_2| \le |z_1| + |z_2| \quad \text{for all } z_1, z_2 \in \mathbb{C}.$$

You may prove this as an exercise!

Also, note that

$$\max\{|\Re(z)|, |\Im(z)|\} \ \le \ |z| \ \le \ |\Re(z)| + |\Im(z)| \quad \text{for all } z \in \mathbb{C}.$$

We shall use complex numbers as scalars and consider matrices whose entries are complex numbers.

The set of all $m \times n$ matrices with entries in $\mathbb{C}$ is denoted by $\mathbb{C}^{m\times n}$. In particular, $\mathbb{C}^{1\times n}$ is the set of all row vectors of length n, while $\mathbb{C}^{m\times 1}$ is the set of all column vectors of length m. For $\mathbf{A} := [a_{jk}] \in \mathbb{C}^{m\times n}$, define $\mathbf{A}^* := [\overline{a_{kj}}]$. Then $\mathbf{A}^* \in \mathbb{C}^{n\times m}$. It is called the **conjugate transpose** or the **adjoint** of $\mathbf{A}$. Note: $(\alpha\mathbf{A} + \beta\mathbf{B})^* = \overline{\alpha}\mathbf{A}^* + \overline{\beta}\mathbf{B}^*$ for $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$ and $\alpha, \beta \in \mathbb{C}$. In case $m = n$, then $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$.

- A square matrix $\mathbf{A} = [a_{jk}]$ is called **Hermitian** or **self-adjoint** if $\mathbf{A}^* = \mathbf{A}$, that is, if $a_{jk} = \overline{a_{kj}}$ for all j, k.
- A square matrix $\mathbf{A} = [a_{jk}]$ is called **skew-Hermitian** or **skew self-adjoint** if $a_{jk} = -\overline{a}_{kj}$ for all j, k.

Note: Every diagonal entry of a self-adjoint matrix is real since $a_{jj} = \overline{a_{jj}} \implies a_{jj} \in \mathbb{R}$ for $j = 1, \ldots, n$. On the other hand, the real part of every diagonal entry of a skew self-adjoint matrix is equal to 0 since $a_{jj} = -\overline{a_{jj}} \implies \Re(a_{jj}) = 0$ for $j = 1, \ldots, n$.

Note: If $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{C}^{n\times 1}$ is a column vector, then $\mathbf{x}^* = \begin{bmatrix} \overline{x_1} & \cdots & \overline{x_n} \end{bmatrix} \in \mathbb{C}^{1\times n}$ is a row vector, and $\mathbf{x}^*\mathbf{x} = |x_1|^2 + \cdots + |x_n|^2$. It follows that $\mathbf{x}^*\mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

A matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ defines a linear transformation from $\mathbb{C}^{n\times 1}$ to $\mathbb{C}^{m\times 1}$, and every linear transformation from $\mathbb{C}^{n\times 1}$ to $\mathbb{C}^{m\times 1}$ can be represented by a matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ (with respect to an ordered basis for $\mathbb{C}^{n\times 1}$ and an ordered basis for $\mathbb{C}^{m\times 1}$).

Similarly, we can consider vector subspaces of $\mathbb{C}^{n\times 1}$, and the concepts of linear dependence of vectors and of the span of a subset carry over to $\mathbb{C}^{n\times 1}$. The Fundamental Theorem for Linear Systems remains valid for matrices with complex entries.

Having thus completed our discussion of solution of a linear system, we shall turn to solution of an 'eigenvalue problem' associated with a matrix. In this development, the role of complex numbers will turn out to be important.

Matrix Eigenvalue Problem

The German word ‘eigen’ means ‘belonging to itself’. The eigenvalue problem for a matrix consists of finding a nonzero vector which is sent to a scalar multiple of itself by the linear transformation defined by the matrix.

Eigenvalue problems come up frequently in many engineering branches, quantum mechanics, physical chemistry, biology, and even in economics and psychology.

Please refer to Section 8.2 of Kreyszig’s book for applications of eigenvalue problems to stretching of elastic membranes, to vibrating mass-spring systems, to Markov processes and to population control models.

In the development that follows, we shall use either real numbers or complex numbers as scalars. To facilitate a general discussion which applies to both types of scalars, we shall write $\mathbb{K}$ to mean either $\mathbb{R}$ or $\mathbb{C}$. When we want to switch to a special treatment valid for only the real scalars, or only for the complex scalars, we shall specify $\mathbb{K} := \mathbb{R}$ or $\mathbb{K} := \mathbb{C}$.

Definition

Let $\mathbf{A}$ be an $n \times n$ matrix with entries in $\mathbb{K}$, that is, let $\mathbf{A} \in \mathbb{K}^{n \times n}$. A scalar $\lambda \in \mathbb{K}$ is called an **eigenvalue** *of $\mathbf{A}$ if*

$$\mathbf{A}\mathbf{x} = \lambda\,\mathbf{x} \quad \textit{for some } \mathbf{x} \in \mathbb{K}^{n \times 1} \textit{ with } \mathbf{x} \neq \mathbf{0}.$$

Any nonzero vector $\mathbf{x} \in \mathbb{K}^{n \times 1}$ satisfying $\mathbf{A}\mathbf{x} = \lambda\,\mathbf{x}$ is called an **eigenvector** *of $\mathbf{A}$ corresponding to the eigenvalue λ. Further,*

$$\{\mathbf{x} \in \mathbb{K}^{n \times 1} : \mathbf{A}\mathbf{x} = \lambda\,\mathbf{x}\} = \mathcal{N}(\mathbf{A} - \lambda\mathbf{I}).$$

is called the **eigenspace** *of $\mathbf{A}$ associated with λ.*

How to find eigenvalues

Let $\mathbf{A} = [a_{jk}] \in \mathbb{K}^{n \times n}$ and let $\lambda \in \mathbb{K}$. Clearly,

$$\begin{aligned} \lambda \text{ is an eigenvalue of } \mathbf{A} &\iff \mathcal{N}(\mathbf{A} - \lambda \mathbf{I}) \neq \{\mathbf{0}\} \\ &\iff \operatorname{rank}(\mathbf{A} - \lambda \mathbf{I}) < n \\ &\iff \det(\mathbf{A} - \lambda \mathbf{I}) = 0. \end{aligned}$$

The last condition suggests that we consider the polynomial

$$p_{\mathbf{A}}(t) := \det(\mathbf{A} - t\,\mathbf{I}) = \det \begin{bmatrix} a_{11} - t & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} - t \end{bmatrix}.$$

This is called the **characteristic polynomial** of $\mathbf{A}$. It is a polynomial of degree n with coefficients in $\mathbb{K}$ and for $\lambda \in \mathbb{K}$,

$$\lambda \text{ is an eigenvalue of } \mathbf{A} \iff \lambda \text{ is a root of } p_{\mathbf{A}}, \text{ i.e., } p_{\mathbf{A}}(\lambda) = 0.$$

In particular, an $n \times n$ matrix with entries in $\mathbb{K}$ has at most n eigenvalues in $\mathbb{K}$.

Algebraic and Geometric Multiplicities

Definition

Let $\mathbf{A} = [a_{jk}] \in \mathbb{K}^{n \times n}$ *and let* $\lambda \in \mathbb{K}$ *be an eigenvalue of* $\mathbf{A}$*.*

- *The* **algebraic multiplicity** *of* λ *(as an eigenvalue of* $\mathbf{A}$*) is the order* m *of the root* λ *of* $p_{\mathbf{A}}(t)$*, i.e.,* m *is the largest positive integer such that* $(t - \lambda)^m$ *divides* $p_{\mathbf{A}}(t)$*.*
- *The* **geometric multiplicity** *of* λ *(as an eigenvalue of* $\mathbf{A}$*) is the dimension of its eigenspace, i.e.,* $\dim \mathcal{N}(\mathbf{A} - \lambda \mathbf{I})$*.*

Observe that if $\lambda \in \mathbb{K}$ be an eigenvalue of $\mathbf{A}$, then

geometric multiplicity of $\lambda = \text{nullity}(\mathbf{A} - \lambda \mathbf{I}) = n - \text{rank}(\mathbf{A} - \lambda \mathbf{I})$.

This can be calculated by solving the homogeneous system $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$ using, for instance, Gaussian elimination. In fact, GEM and back substitution will also give the *basic solutions*, i.e., a set of eigenvectors of $\mathbf{A}$ which forms a basis of the eigenspace of $\mathbf{A}$ associated to λ.

Examples

(i) Let $\mathbf{A} = \text{diag}(a_1, \ldots, a_n)$, i.e., let $\mathbf{A}$ be a diagonal matrix with diagonal entries $a_1, \ldots, a_n$ in that order. Clearly

$$p_{\mathbf{A}}(t) = (a_1 - t)(a_2 - t) \cdots (a_n - t).$$

Thus the eigenvalues of $\mathbf{A}$ are precisely $a_1, \ldots, a_n$. Note that this can also be seen directly since $\mathbf{A}\mathbf{e}_k = a_k \mathbf{e}_k$ for each $k = 1, \ldots, n$, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the basic column vectors in $\mathbb{K}^{n \times 1}$. Observe that in this case the algebraic multiplicity of each eigenvalue is equal to its geometric multiplicity. Indeed, if $\lambda \in \{a_1, \ldots, a_n\}$, then the algebraic multiplicity of λ equals

$$m := \text{the number of } i \in \{1, \ldots, n\} \text{ such that } a_i = \lambda.$$

To find the geometric multiplicity of λ, consider $\mathbf{A} - \lambda\mathbf{I}$ and note that this is a diagonal matrix with exactly m rows of zeros and $n - m$ nonzero rows. So $\text{rank}(\mathbf{A} - \lambda\mathbf{I}) = n - m$ and hence

$$\text{geometric multiplicity of } \lambda = \text{nullity}(\mathbf{A} - \lambda\mathbf{I}) = m.$$

Examples Contd.

(ii) Let $\mathbf{A}$ be an upper triangular matrix with diagonal entries $a_{11}, \ldots, a_{nn}$. Again, the characteristic polynomials factors as

$$p_{\mathbf{A}}(t) = (a_{11} - t)(a_{22} - t) \cdots (a_{nn} - t).$$

So the eigenvalues of $\mathbf{A}$ are precisely $a_{11}, \ldots, a_{nn}$. The algebraic multiplicities can be found as in the previous example. However, they may not always coincide with the corresponding geometric multipliciities.

For example, consider the 2×2 matrix $\mathbf{A} := \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$. Clearly 3 is the only eigenvalue of $\mathbf{A}$ and its algebraic multiplicity is 2. On the other hand, the homogeneous linear system $(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{0}$ comprises of the equations $x_2 = 0$ and $0 = 0$. So the eigenspace of $\mathbf{A}$ associated with the eigenvalue 3 has $[1\ 0]^{\mathsf{T}}$ as its basis. Thus the geometric multiplicity of the eigenvalue 3 of $\mathbf{A}$ is 1.

Remarks on finding eigenvalues and eigenvectors

In general, solving an eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{x}$ is much harder than finding solutions of a linear system $\mathbf{Ax} = \mathbf{b}$. In the latter case, the matrix $\mathbf{A}$ and the right side $\mathbf{b}$ are given. On the other hand, in the eigenvalue problem, the 'unknown' vector $\mathbf{x}$ appears on both sides of the equation, and additionally, an 'unknown' scalar λ appears on the right side.

We need to find an eigenvalue λ of $\mathbf{A}$ and a corresponding eigenvector $\mathbf{x}$ of $\mathbf{A}$ simultaneously. It is tough, but if one of them is known beforehand, then the other can be found easily. Suppose a scalar λ is known to be an eigenvalue of $\mathbf{A}$. Then all eigenvectors of $\mathbf{A}$ corresponding to λ can be obtained by finding the general solution of the homogeneous linear system $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$. Next, suppose a nonzero vector $\mathbf{x}$ is known to be an eigenvector of $\mathbf{A}$. Then one only needs to calculate $\mathbf{Ax}$ and observe that it is a scalar multiple of $\mathbf{x}$. This scalar is the corresponding eigenvalue of $\mathbf{A}$.

Examples

(i) Let $\mathbf{A} := \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$, and suppose somehow we know that $\lambda := -3$ is an eigenvalue of $\mathbf{A}$. Let

$$\mathbf{B} := \mathbf{A} - (-3)\mathbf{I} = \mathbf{A} + 3\mathbf{I} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix}.$$

By EROs, we can transform $\mathbf{B}$ to $\mathbf{B}' := \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Now the solution space of $\mathbf{B}'\mathbf{x} = \mathbf{0}$ is $\{\mathbf{x} \in \mathbb{R}^{3\times 1} : x_1 + 2x_2 - 3x_3 = 0\}$, which is also the solution space of $\mathbf{B}\mathbf{x} := (\mathbf{A} + 3\mathbf{I})\mathbf{x} = \mathbf{0}$, the basic solutions being $\mathbf{s}_2 := \begin{bmatrix} -2 & 1 & 0 \end{bmatrix}^{\mathsf{T}}$ and $\mathbf{s}_3 := \begin{bmatrix} 3 & 0 & 1 \end{bmatrix}^{\mathsf{T}}$. Thus the eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue $\lambda = -3$ are the nonzero linear combinations of $\mathbf{s}_2$ and $\mathbf{s}_3$. The geometric multiplicity of the eigenvalue -3 is equal to 2.

(ii) Let $\mathbf{A} := \frac{1}{10}\begin{bmatrix} 29 & 6 & -1 \\ 29 & 16 & -11 \\ 25 & 10 & 15 \end{bmatrix}$, and suppose somehow we know that $\mathbf{x} := \begin{bmatrix} 1 & -1 & 3 \end{bmatrix}^{\mathsf{T}}$ is an eigenvector of $\mathbf{A}$. We easily find that $\mathbf{A}\mathbf{x} = 2\begin{bmatrix} 1 & -1 & 3 \end{bmatrix}^{\mathsf{T}}$. Hence is the corresponding eigenvalue of $\mathbf{A}$.

We saw that eigenvalue problems for diagonal matrices are the easiest to solve. We wonder when a nondiagonal matrix would 'behave like a diagonal matrix'. To make this precise, we introduce the following notion.

Definition

Let $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n\times n}$. We say that $\mathbf{A}$ is **similar** *to $\mathbf{B}$ (over $\mathbb{K}$) if there is an invertible $\mathbf{P} \in \mathbb{K}^{n\times n}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, that is, $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{B}$. In this case, we write $\mathbf{A} \sim \mathbf{B}$.*

One can easily check (i) $\mathbf{A} \sim \mathbf{A}$, (ii) if $\mathbf{A} \sim \mathbf{B}$ then $\mathbf{B} \sim \mathbf{A}$, and (iii) if $\mathbf{A} \sim \mathbf{B}$ and $\mathbf{B} \sim \mathbf{C}$, then $\mathbf{A} \sim \mathbf{C}$.

Examples

(i) Let $\mathbf{A} := \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$. Let $\mathbf{P} := \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$, so that $\mathbf{P}^{-1} := \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$. Then $\mathbf{A}$ is similar to the matrix

$$\mathbf{B} := \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 6 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},$$

which is a diagonal matrix.

(ii) Let $\mathbf{A} \in \mathbb{K}^{n \times n}$. Then $\mathbf{A} \sim \mathbf{I} \iff \mathbf{A} = \mathbf{I}$.

(iii) Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let $\mathbf{E}$ be $n \times n$ an elementary matrix. Then $\mathbf{B} := \mathbf{E}\mathbf{A}\mathbf{E}^{-1}$ is similar to $\mathbf{A}$. Note: $\mathbf{EA}$ is obtained from $\mathbf{A}$ by an elementary row operation on $\mathbf{A}$, and $\mathbf{B}$ is obtained from $\mathbf{EA}$ by the 'reverse column operation' on $\mathbf{EA}$.

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 $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{AP}$. Writing $\mathbf{I} = \mathbf{P}^{-1}\mathbf{P}$, we see that

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

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 eigevector of $\mathbf{A}'$ corresponding to λ. Also, it is easy to check that $\{\mathbf{x}_1, \ldots, \mathbf{x}_g\}$ is a basis for $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})$ if and only if $\{\mathbf{P}^{-1}\mathbf{x}_1, \ldots, \mathbf{P}^{-1}\mathbf{x}_g\}$ 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

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

Note that in the above example, if $\mathbf{A}$ were similar to a diagonal matrix $\mathbf{D}$, then we must have $\mathbf{D} = \mathrm{diag}(\lambda, \lambda)$, 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,

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

So $\mathrm{nullity}(\mathbf{A} - 3\mathbf{I}) = 2$. In fact, $\left\{ \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^{\mathsf{T}}, \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}^{\mathsf{T}} \right\}$ 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 $T_{\mathbf{A}} : \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ 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 $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ 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 kth column of LHS is $\mathbf{A}\mathbf{x}_k$ and the kth column of RHS is the linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with coefficients from the kth column of $\mathbf{B}$. Thus $\mathbf{A}\mathbf{x}_k = b_{1k}\mathbf{x}_1 + \cdots + b_{nk}\mathbf{x}_n$ for $k = 1, \ldots, n$. This means the kth column of $\mathbf{M}_E^E(T_{\mathbf{A}})$ is the kth column $\begin{bmatrix} b_{1k} & \cdots & b_{nk} \end{bmatrix}^{\mathsf{T}}$ of $\mathbf{B}$, $k = 1, \ldots, n$, that is, $\mathbf{B} = \mathbf{M}_E^E(T_{\mathbf{A}})$. $\square$

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 $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ 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* $\mathbb{K}$*).*

Proposition (A Characterization of Diagonalizability)

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, $\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{D}$, where $\mathbf{X} := \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}$ and $\mathbf{D} := \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \iff \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a basis for $\mathbb{K}^{n\times 1}$ and $\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$ for $k = 1, \ldots, n$.

Proof. $\mathbf{A}$ is diagonalizable $\iff$ there is an invertible matrix $\mathbf{X}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{D}$. This is the case if and only if there is a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ for $\mathbb{K}^{n\times 1}$ and there are $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ such that

$$\mathbf{A}\begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix} = \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}\mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

The kth column of LHS is $\mathbf{A}\mathbf{x}_k$ and the kth column of RHS is the linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with coefficients $0. \ldots, 0, \lambda_k, 0, \ldots, 0$, that is, $\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$, $k = 1, \ldots, n$. $\square$