MA 110
Linear Algebra and Differential Equations
Lecture 15

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

Spring 2025

Recall that we discussed the following in the last lecture.

- Notion of orthogonality of $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$)
- Orthogonal set $E \subset \mathbb{K}^n$; the subspace $E^\perp$
- Projection of $\mathbf{x}$ in the direction of a nonzero vector $\mathbf{y}$:.

$$P_{\mathbf{y}}(\mathbf{x}) := \frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \mathbf{y}.$$

- Important property: $\big(\mathbf{x} - P_{\mathbf{y}}(\mathbf{x})\big) \perp \mathbf{y}$.
- E orthogonal and $\mathbf{0} \notin E \implies E$ is linearly independent.
- Gram-Schmidt Orthogonalization Process (G-S OP)
- Notion of an orthonormal set.
- Existence of an orthonormal basis for every nonzero vector subspace
- Analogue for row vectors

After a detour of inner products and orthonormal sets, we come back to the matrix eigenvalue problem. We shall show that if the scalars are complex numbers, then every square matrix $\mathbf{A}$ can be 'upper triangularized', that is, it is similar to an upper triangular matrix $\mathbf{B} := \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. In fact, we shall show that $\mathbf{P}$ can be chosen to be of a particularly nice matrix

Definition

A matrix $\mathbf{U} \in \mathbb{K}^{n \times n}$ *is called* **unitary** *if the columns of* $\mathbf{U}$ *form an orthonormal subset of* $\mathbb{K}^{n \times 1}$*. In that case, the columns of* $\mathbf{U}$ *are, in fact, an orthonormal basis for* $\mathbb{K}^{n \times 1}$*.*

Proposition

A matrix is unitary if and only if it is invertible and its inverse is the same as its adjoint.

Proof. Let $\mathbf{U} \in \mathbb{K}^{n \times n}$ be unitary. Then $\operatorname{rank} \mathbf{U} = n$ since the n columns $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of $\mathbf{U}$ form a basis for $\mathbb{K}^{n \times 1}$. Hence $\mathbf{U}$ is invertible. Further, because of the orthonormality,

$$\mathbf{U}^*\mathbf{U} = \begin{bmatrix} \mathbf{u}_1^* \\ \vdots \\ \mathbf{u}_n^* \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^*\mathbf{u}_1 & \cdots & \mathbf{u}_1^*\mathbf{u}_n \\ \vdots & \vdots & \vdots \\ \mathbf{u}_n^*\mathbf{u}_1 & \cdots & \mathbf{u}_n^*\mathbf{u}_n \end{bmatrix} = \mathbf{I}.$$

It follows that $\mathbf{U}\mathbf{U}^* = \mathbf{I}$ as well. Hence $\mathbf{U}^{-1} = \mathbf{U}^*$.

Conversely, the above calculation shows that if a square matrix $\mathbf{U}$ satisfies $\mathbf{U}^*\mathbf{U} = \mathbf{I}$, then its columns form an orthonormal subset of $\mathbb{K}^{n\times 1}$, that is, $\mathbf{U}$ is a unitary matrix. □

We note that if $\mathbf{A}$ and $\mathbf{B}$ are unitary, then so is $\mathbf{AB}$ since $(\mathbf{AB})^*(\mathbf{AB}) = (\mathbf{B}^*\mathbf{A}^*)(\mathbf{AB}) = \mathbf{B}(\mathbf{A}^*\mathbf{A})\mathbf{B} = \mathbf{B}^*\mathbf{B} = \mathbf{I}$.

Examples

(i) The $n\times n$ identity matrix $\mathbf{I}_n := \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \end{bmatrix}$ is unitary.

(ii) For $\theta \in \mathbb{R}$, the matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$ is unitary.

It represents a rotation about the x_1-axis in $\mathbb{R}^{3\times 1}$.

(iii) The matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ is unitary. It represents reflection about the x_1x_2-plane in $\mathbb{R}^{3\times 1}$.

(iv) If $\mathbb{K} = \mathbb{C}$, then the matrix $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -i/\sqrt{2} & i/\sqrt{2} & 0 \\ 0 & 0 & i \end{bmatrix}$ is unitary.

(v) The matrix $\begin{bmatrix} 1/\sqrt{3} & 4/\sqrt{42} & -4/\sqrt{105} & 2/\sqrt{30} \\ 1/\sqrt{3} & -5/\sqrt{42} & -2/\sqrt{105} & 1/\sqrt{30} \\ 0 & 0 & -7/\sqrt{105} & -4/\sqrt{30} \\ 1/\sqrt{3} & 1/\sqrt{42} & 6/\sqrt{105} & -3/\sqrt{30} \end{bmatrix}$,

whose columns are obtained by orthonormalizing the column vectors $\begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix}^{\mathsf{T}}$, $\begin{bmatrix} 1 & -2 & 0 & 0 \end{bmatrix}^{\mathsf{T}}$, $\begin{bmatrix} 1 & 0 & -1 & 2 \end{bmatrix}^{\mathsf{T}}$ and $\begin{bmatrix} 2 & 1 & -4 & -3 \end{bmatrix}^{\mathsf{T}}$ in $\mathbb{K}^{4\times 1}$ is unitary.

Definition

Let $\mathbf{A}$, $\mathbf{B} \in \mathbb{K}^{n\times n}$. *We say that* $\mathbf{A}$ *is* **unitarily similar** *to* $\mathbf{B}$ *if there is a unitary matrix* $\mathbf{U}$ *such that* $\mathbf{B} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. *Also,* $\mathbf{A}$ *is called* **unitarily diagonalizable** *if it is unitarily similar to a diagonal matrix.*

The following result is analogous to the result proved earlier that related similarity of matrices with change of basis.

Proposition

Let $\mathbf{A}$, $\mathbf{B} \in \mathbb{K}^{n\times n}$. Then $\mathbf{A}$ is unitarily similar to $\mathbf{B}$ if and only if there is an ordered orthonormal 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, $\mathbf{B} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$ if and only if the columns of $\mathbf{U}$ form an ordered orthonormal 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 $\mathbf{A}$ is unitarily similar to $\mathbf{B}$ if and only if there is a unitary matrix $\mathbf{U}$ such that $\mathbf{AU} = \mathbf{UB}$. This is the case if and only if there is an ordered orthonormal basis $E := (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ for $\mathbb{K}^{n\times 1}$ such that

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

Comparing the kth columns of the LHS and RHS, we obtain

$$\mathbf{A}\mathbf{u}_k = b_{1k}\mathbf{u}_1 + \cdots + b_{nk}\mathbf{u}_n \quad \text{for } k = 1, \ldots, n.$$

This means that 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 $\mathbf{B}$ for $k = 1, \ldots, n$, i.e., $\mathbf{B} = \mathbf{M}_E^E(T_\mathbf{A})$. □

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 $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ for $\mathbb{K}^{n\times 1}$, the matrix $\mathbf{B} := \mathbf{U}^{-1}\mathbf{AU}$ is the matrix of the same linear transformation $T_\mathbf{A}$ with respect to the ordered orthonormal basis for $\mathbb{K}^{n\times 1}$ consisting of the columns of $\mathbf{U}$.

By specializing the proof of the above proposition to the case when $\mathbf{B}$ is the diagonal matrix $\mathbf{D} := \text{diag}(\lambda_1, \ldots, \lambda_n)$, and noting that in this case, the kth column of $\begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \mathbf{D}$ is simply $\lambda_k \mathbf{u}_k$, we obtain the following result.

Proposition

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

$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \mathbf{D}$ for a unitary matrix $\mathbf{U} := \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}$ and for $\mathbf{D} := \text{diag}(\lambda_1, \ldots, \lambda_n) \iff \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthonormal basis for $\mathbb{K}^{n \times 1}$ and $\mathbf{A}\mathbf{u}_k = \lambda_k \mathbf{u}_k$ for $k = 1, \ldots, n$.

Remark: Unitary matrices over $\mathbb{R}$ are often called orthogonal matrices. Thus, when $\mathbb{K} = \mathbb{R}$, we speak of a matrix being orthogonally diagonalized rather than unitarily diagonalized.

Question: Can we find intrinsic conditions on a matrix that characterise when it can be unitarily diagonalized?

Eigenvalues and eigenvectors of linear maps

The definition is analogous to that for matrices.

Definition

Let V be a vector subspace of dimension n (possibly contained in a higher dimensional space of column vectors). Let $T : V \to V$ be a linear transformation. We say that $\lambda \in \mathbb{K}$ is an **eigenvalue** *of T if there is nonzero $\mathbf{x} \in V$ such that $T(\mathbf{x}) = \lambda \mathbf{x}$; such a nonzero vector is called an* **eigenvector** *of T corresponding to λ.*

Recall that if $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is an ordered basis for V, then the linear transformation T can be represented by the $n \times n$ matrix $\mathbf{A} := \mathbf{M}_E^E(T)$ whose kth column is $\begin{bmatrix} a_{1k} & \cdots & a_{nk} \end{bmatrix}^{\mathsf{T}}$, where $T(\mathbf{x}_k) = a_{1k}\mathbf{x}_1 + \cdots + a_{nk}\mathbf{x}_n$.

Lemma

Let V be an n dimensional vector subspace, and let E be an ordered basis for V. Suppose $T : V \to V$ is a linear map. Then $\lambda \in \mathbb{K}$ is an eigenvalue of T if and only if λ is an eigenvalue of $M_E^E(T)$.
In particular, if $\mathbb{K} := \mathbb{C}$, then T has an eigenvalue.

Proof. Let $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $\mathbf{A} := M_E^E(T) = [a_{jk}]$.
Consider $\lambda \in \mathbb{K}$. Let $\mathbf{x} := \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n \in V$. Then

$$T(\mathbf{x}) = \lambda\,\mathbf{x} \iff \mathbf{A}\begin{bmatrix}\alpha_1 & \cdots & \alpha_n\end{bmatrix}^{\mathsf{T}} = \lambda\begin{bmatrix}\alpha_1 & \cdots & \alpha_n\end{bmatrix}^{\mathsf{T}}.$$

Also, $\mathbf{x} \neq \mathbf{0} \iff \begin{bmatrix}\alpha_1 & \cdots & \alpha_n\end{bmatrix}^{\mathsf{T}} \neq \mathbf{0}$. Hence $\mathbf{x}$ is an eigenvector of $T \iff \begin{bmatrix}\alpha_1 & \cdots & \alpha_n\end{bmatrix}^{\mathsf{T}}$ is an eigenvector of $\mathbf{A}$.

Let $\mathbb{K} = \mathbb{C}$. Since $\mathbf{A}$ has an eigenvalue, so does T □