MA 110
Linear Algebra and Differential Equations
Lecture 17

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.

- Schur's Theorem: Any $\mathbf{A} \in \mathbb{C}^{n\times n}$ is unitarily similar to an upper triangular matrix.
- Trace of $\mathbf{A} \in \mathbb{C}^{n\times n}$ as sum of eigevalues, and det $\mathbf{A}$ as product of eigenvalues. Also polynomial equations satisfied by $\mathbf{A}$ are satisfied by every eigenvalue of $\mathbf{A}$. Examples.
- $\mathbf{A} \in \mathbb{K}^{n\times n}$ is a normal matrix if $\mathbf{A}^*\mathbf{A} = \mathbf{A}\,\mathbf{A}^*$.
- If $\mathbf{A} \in \mathbb{C}^{n\times n}$ is self-adjoint ($\mathbf{A}^* = \mathbf{A}$), or skew self-adjoint ($\mathbf{A}^* = -\mathbf{A}$), or unitary ($\mathbf{A}^* = \mathbf{A}^{-1}$), then $\mathbf{A}$ is normal. In particular, if $\mathbf{A} \in \mathbb{R}^{n\times n}$ is symmetric, skew-symmetric, or orthogonal, then $\mathbf{A}$ is normal. But if $\mathbf{A}$ is normal, then it need not have any of the above properties. Also, there exist $n \times n$ matrices over $\mathbb{K}$ that are not normal.

Proposition

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$. Then $\mathbf{A}$ is normal if and only if

$$\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle \quad \textit{for all } \mathbf{x}, \mathbf{y} \in \mathbb{K}^{n \times 1}.$$

Proof. For $\mathbf{x}, \mathbf{y}$, $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = (\mathbf{A}\mathbf{x})^*\mathbf{A}\mathbf{y} = \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{y}$ and $\langle \mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle = (\mathbf{A}^*\mathbf{x})^*\mathbf{A}^*\mathbf{y} = \mathbf{x}^*\mathbf{A}\,\mathbf{A}^*\mathbf{y}$.

If $\mathbf{A}$ is normal, then clearly, $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n \times 1}$. Conversely, suppose $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n \times 1}$. Then letting $\mathbf{x} := \mathbf{e}_j$ and $\mathbf{y} := \mathbf{e}_k$, we see that $\mathbf{e}_j^*\mathbf{A}^*\mathbf{A}\,\mathbf{e}_k = \mathbf{e}_j^*\mathbf{A}\,\mathbf{A}^*\mathbf{e}_k$, that is, the (j, k)th entries of $\mathbf{A}^*\mathbf{A}$ and $\mathbf{A}\,\mathbf{A}^*$ are the same for all $j, k = 1, \ldots, n$. Thus we see that $\mathbf{A}^*\mathbf{A} = \mathbf{A}\,\mathbf{A}^*$, that is, $\mathbf{A}$ is normal. □

The above condition for normality of a matrix $\mathbf{A}$ says that the lengths of the vectors $\mathbf{A}\mathbf{x}$ and $\mathbf{A}^*\mathbf{x}$ should be the same, and the angle between $\mathbf{A}\mathbf{x}$ and $\mathbf{A}\mathbf{y}$ should be the same as the angle between $\mathbf{A}^*\mathbf{x}$ and $\mathbf{A}^*\mathbf{y}$ for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{K}^{n \times 1}$.

Corollary

Let $\mathbf{A} \in \mathbb{K}^{n\times n}$ be normal. Then $\|\mathbf{A}\mathbf{x}\| = \|\mathbf{A}^*\mathbf{x}\|$ for all $\mathbf{x}$ in $\mathbb{K}^{n\times 1}$. Let $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ corresponding to an eigenvalue λ. Then $\mathbf{x}$ itself is an eigenvector of $\mathbf{A}^*$ corresponding to the eigenvalue $\overline{\lambda}$ of $\mathbf{A}^*$. Further, if $\mathbf{y}$ is an eigenvector of $\mathbf{A}$ corresponding to an eigenvalue $\mu \neq \lambda$, then $\mathbf{y}$ is orthogonal to $\mathbf{x}$.

Proof. Let $\mathbf{x} \in \mathbb{K}^{n\times 1}$. Since $\mathbf{A}$ is normal,

$$\|\mathbf{A}\mathbf{x}\|^2 = \langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{x}\rangle = \langle \mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{x}\rangle = \|\mathbf{A}^*\mathbf{x}\|^2 \quad \text{for all } \mathbf{x} \in \mathbb{K}^{n\times 1}.$$

Next, let $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ corresponding to an eigenvalue λ. Then $\|\mathbf{A}^*\mathbf{x} - \overline{\lambda}\mathbf{x}\| = \|\mathbf{A}\mathbf{x} - \lambda\mathbf{x}\| = 0$. Hence $\mathbf{x}$ itself is an eigenvector $\mathbf{A}^*$ corresponding to the eigenvalue $\overline{\lambda}$. Finally, let $\mathbf{y}$ be an eigenvector of $\mathbf{A}$ corresponding to an eigenvalue $\mu \neq \lambda$. Then

$$\mu\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, \mu\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{A}\mathbf{y}\rangle = \langle \mathbf{A}^*\mathbf{x}, \mathbf{y}\rangle = \langle \overline{\lambda}\mathbf{x}, \mathbf{y}\rangle = \lambda\langle \mathbf{x}, \mathbf{y}\rangle.$$

Since $\mu \neq \lambda$, we see that $\langle \mathbf{x}, \mathbf{y}\rangle = 0$, that is, $\mathbf{x} \perp \mathbf{y}$. □

We now characterize a diagonal matrix in terms of its upper triangularity and normality.

Lemma

A matrix $\mathbf{A} \in \mathbb{K}^{n\times n}$ is diagonal if and only if it is upper triangular and normal.

Proof. Let $\mathbf{A} := \text{diag}(\lambda_1, \ldots, \lambda_n)$. Then $\mathbf{A}$ is upper triangular. Also, it is normal since $\mathbf{A}^*\mathbf{A} = \text{diag}(|\lambda_1|^2, \ldots, |\lambda_n|^2) = \mathbf{A}\,\mathbf{A}^*$.

Conversely, suppose $\mathbf{A} := [a_{jk}]$ is upper triangular and normal. Let $\mathbf{B} := \mathbf{A}^*\mathbf{A} = [b_{jk}]$ and $\mathbf{C} := \mathbf{A}\mathbf{A}^* = [c_{jk}]$. Since $\mathbf{A}$ is upper triangular, $a_{jk} = 0$ if $j > k$, and so for $k = 1, \ldots, n$,

$$b_{kk} = \sum_{\ell=1}^{n} \overline{a}_{\ell k} a_{\ell k} = \sum_{\ell=1}^{k} |a_{\ell k}|^2 \quad \text{and} \quad c_{kk} = \sum_{\ell=1}^{n} a_{k\ell}\overline{a}_{k\ell} = \sum_{\ell=k}^{n} |a_{k\ell}|^2.$$

Since $\mathbf{A}$ is normal, we see that $b_{kk} = c_{kk}$ for $k = 1, \ldots, n$.

Let $k = 1$. Then
$|a_{11}|^2 = b_{11} = c_{11} = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2$.
Hence $\boxed{a_{12} = \cdots = a_{1n} = 0}$.

Next, let $k = 2$. Then
$|a_{22}|^2 = |a_{12}|^2 + |a_{22}|^2 = b_{22} = c_{22} = |a_{22}|^2 + |a_{23}|^2 + \cdots + |a_{2n}|^2$.
Hence $\boxed{a_{23} = \cdots = a_{2n} = 0}$.

Proceeding in this manner, for $k = n - 1$, we obtain
$|a_{(n-1)(n-1)}|^2 = |a_{1(n-1)}|^2 + \cdots + |a_{(n-1)(n-1)}|^2 = b_{(n-1)(n-1)} = c_{(n-1)(n-1)} = |a_{(n-1)(n-1)}|^2 + |a_{(n-1)n}|^2$.
Hence $\boxed{a_{(n-1)n} = 0}$.

Thus $a_{jk} = 0$ if $j < k$, that is, $\mathbf{A}$ is lower triangular. Since $\mathbf{A}$ is given to be upper triangular, it is is in fact diagonal. □

We are now ready to state and prove necessary conditions as well as sufficient conditions for diagonalizing a matrix unitarily.

Proposition (Spectral Theorem for Normal Matrices)

(i) If $\mathbf{A} \in \mathbb{K}^{n\times n}$ is unitarily diagonalizable, then $\mathbf{A}$ is normal.

(ii) If $\mathbf{A} \in \mathbb{C}^{n\times n}$ is normal, then $\mathbf{A}$ is unitarily diagonalizable.

Proof.

(i) Let $\mathbf{U}$ be a unitary matrix and $\mathbf{D}$ be a diagonal matrix such that $\mathbf{D} = \mathbf{U}^*\mathbf{A}\mathbf{U}$. Then $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$, and so $\mathbf{A}^* = \mathbf{U}\mathbf{D}^*\mathbf{U}^*$. Further, since $\mathbf{D}$ is always normal, $\mathbf{D}^*\mathbf{D} = \mathbf{D}\mathbf{D}^*$. Hence

$$\begin{aligned}\mathbf{A}^*\mathbf{A} &= (\mathbf{U}\mathbf{D}^*\mathbf{U}^*)(\mathbf{U}\mathbf{D}\mathbf{U}^*) = \mathbf{U}(\mathbf{D}^*\mathbf{D})\mathbf{U}^* \\ &= \mathbf{U}(\mathbf{D}\mathbf{D}^*)\mathbf{U}^* = (\mathbf{U}\mathbf{D}\mathbf{U}^*)(\mathbf{U}\mathbf{D}^*\mathbf{U}^*) \\ &= \mathbf{A}\,\mathbf{A}^*.\end{aligned}$$

Thus $\mathbf{A}$ is normal.

(ii) Let $\mathbf{A} \in \mathbb{C}^{n\times n}$. By Schur's theorem, there is a unitary matrix $\mathbf{U}$ and an upper triangular matrix $\mathbf{B}$ with $\mathbf{B} = \mathbf{U}^*\mathbf{A}\mathbf{U}$.

Suppose $\mathbf{A}$ is normal, that is, $\mathbf{A}^*\mathbf{A} = \mathbf{A}\,\mathbf{A}^*$. Then using $\mathbf{B} = \mathbf{U}^*\mathbf{A}\mathbf{U}$, we can readily check that $\mathbf{B}^*\mathbf{B} = \mathbf{B}\,\mathbf{B}^*$, that is, $\mathbf{B}$ is normal.

Since $\mathbf{B}$ is upper triangular and normal, $\mathbf{B}$ is diagonal by the previous lemma. Thus $\mathbf{A}$ is unitarily diagonalizable. □

Remarks

(i) The unitary matrix $\mathbf{U}$ and the diagonal matrix $\mathbf{B}$ such that $\mathbf{A} = \mathbf{U}\mathbf{B}\mathbf{U}^*$ are not unique. We shall give some examples later.

(ii) Part (ii) of the proposition is not true for real scalars, that is, even if $\mathbf{A} \in \mathbb{R}^{n\times n}$ is normal, there may be no unitary $\mathbf{U} \in \mathbb{R}^{n\times n}$ and diagonal $\mathbf{D} \in \mathbb{R}^{n\times n}$ such that $\mathbf{D} = \mathbf{U}^*\mathbf{A}\mathbf{U}$. For example, $\mathbf{A} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ is normal, but it is not diagonalizable using real scalars since it has no real eigenvalue.

Just as we have proved that a matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ is normal if and only if it is unitarily diagonalizable, we prove a similar result for self-adjoint matrices. We will give a proof that avoids the use of Spectral Theorem for Normal Matrices.

Proposition (Spectral Theorem for Self-Adjoint Matrices)

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$. Then $\mathbf{A}$ is self-adjoint if and only if $\mathbf{A}$ is unitarily diagonalizable and all eigenvalues of $\mathbf{A}$ are real.

Proof. Suppose $\mathbf{A}$ is unitarily diagonalizable and all eigenvalues of $\mathbf{A}$ are real. Let $\mathbf{U}$ be a unitary matrix and let $\mathbf{D}$ be a diagonal matrix such that $\mathbf{D} = \mathbf{U}^*\mathbf{A}\mathbf{U}$. Then the diagonal entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$, and so they are real. Hence $\mathbf{D}^* = \mathbf{D}$. Consequently,

$$\mathbf{A}^* = (\mathbf{U}\mathbf{D}\mathbf{U}^*)^* = \mathbf{U}\mathbf{D}^*\mathbf{U}^* = \mathbf{U}\mathbf{D}\mathbf{U}^* = \mathbf{A}.$$

Thus $\mathbf{A}$ is self-adjoint.

Conversely, suppose $\mathbf{A}$ is self-adjoint. By Schur's theorem, there is a unitary matrix $\mathbf{U}$, and also an upper triangular matrix $\mathbf{B}$ such that $\mathbf{B} = \mathbf{U}^*\mathbf{A}\mathbf{U}$. Then

$$\mathbf{B}^* = (\mathbf{U}^*\mathbf{A}\mathbf{U})^* = \mathbf{U}^*\mathbf{A}^*\mathbf{U} = \mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{B},$$

so that $\mathbf{B}$ is self-adjoint. Since $\mathbf{B}$ is upper triangular and self-adjoint, the previous lemma shows that $\mathbf{B}$ is in fact diagonal with all diagonal entries real. Thus $\mathbf{A}$ is unitarily diagonalizable. Also, all eigenvalues of $\mathbf{A}$ are real, since they are the diagonal entries of the matrix $\mathbf{B}$. □

Our short proof of the spectral theorem for self-adjoint matrices is based on Schur's theorem. This result can also be deduced from part (ii) of the spectral theorem for normal matrices since every self-adjoint matrix in normal, provided we independently show that every eigenvalue of a self-adjoint matrix is real. The latter statement can be easily proved.

Proposition

If $\mathbf{A} \in \mathbb{C}^{n \times n}$ is self-adjoint, then every eigenvalue of $\mathbf{A}$ is real.

Proof.
Let λ be an eigenvalue of a self-adjoint matrix $\mathbf{A}$, and let $\mathbf{x}$ be a corresponding unit eigenvector. Then

$$\lambda = \lambda \, \mathbf{x}^*\mathbf{x} = \mathbf{x}^*\mathbf{A}\,\mathbf{x} = \mathbf{x}^*\mathbf{A}^*\mathbf{x} = (\mathbf{A}\,\mathbf{x})^*\mathbf{x} = (\lambda\,\mathbf{x})^*\mathbf{x} = \overline{\lambda}\,\mathbf{x}^*\mathbf{x} = \overline{\lambda}.$$

Hence λ is real. □

Finally, let us consider a real symmetric matrix $\mathbf{A}$, that is, $\mathbf{A} \in \mathbb{R}^{n \times n}$ such that $\mathbf{A}^{\mathsf{T}} = \mathbf{A}$. We shall prove a spectral theorem for $\mathbf{A}$ which involves only real scalars.

A unitary matrix with real entries is also called an **orthogonal matrix**. Thus $\mathbf{C} \in \mathbb{R}^{n \times n}$ is orthogonal if its columns form an orthonormal subset of $\mathbb{R}^{n \times 1}$. Clearly, an orthogonal matrix is invertible and its inverse is the same as its transpose.

A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is called **orthogonally diagonalizable** if there is an orthogonal matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$ such that $\mathbf{D} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$. We prove Jacobi's theorem.

Proposition (Spectral Theorem for Real Symmetric Matrices)

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$. Then $\mathbf{A}$ is symmetric if and only if $\mathbf{A}$ is orthogonally diagonalizable. In this case, $\mathbf{A}$ has n real eigenvalues counted according to their algebraic multiplicities.

Proof.
Suppose $\mathbf{A}$ is orthogonally diagonalizable. Let $\mathbf{C} \in \mathbb{R}^{n \times n}$ be an orthogonal matrix, and let $\mathbf{D} \in \mathbb{R}^{n \times n}$ be a diagonal matrix such that $\mathbf{D} = \mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C}$. Since $\mathbf{D}^{\mathsf{T}} = \mathbf{D}$,

$$\mathbf{A}^{\mathsf{T}} = (\mathbf{C}\mathbf{D}\mathbf{C}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{C}\mathbf{D}^{\mathsf{T}}\mathbf{C}^{\mathsf{T}} = \mathbf{C}\mathbf{D}\mathbf{C}^{\mathsf{T}} = \mathbf{A}.$$

Thus $\mathbf{A}$ is symmetric.

Conversely, suppose $\mathbf{A}$ is symmetric. Since $\mathbb{R}$ can be considered as a subset of $\mathbb{C}$, we treat $\mathbf{A}$ as a matrix with complex entries. Then $\mathbf{A}^* = \mathbf{A}$, that is, $\mathbf{A}$ is self-adjoint, and so all eigenvalues of $\mathbf{A}$ are real.

By the spectral theorem for self-adjoint matrices, there is a unitary matrix $\mathbf{U} \in \mathbb{C}^{n\times n}$ and a diagonal matrix $\mathbf{D} \in \mathbb{C}^{n\times n}$ such that $\mathbf{D} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$, that is, $\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{D}$. Since the diagonal entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$, we see that $\mathbf{D} \in \mathbb{R}^{n\times n}$.

The n columns of the unitary matrix $\mathbf{U}$ form an orthonormal set in $\mathbb{C}^{n\times n}$, and each column is an eigenvector of $\mathbf{A}$ corresponding to an eigenvalue of $\mathbf{A}$.

Let λ be an eigenvalue of $\mathbf{A}$. Then $\lambda \in \mathbb{R}$. Using the Gauss Elimination Method, we may find the basic solutions of the homogeneous linear system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$. Their entries are real since all entries of $\mathbf{A}$ are real and $\lambda \in \mathbb{R}$. These solutions form a basis for the eigenspace of $\mathbf{A}$ corresponding to λ.

Further, we can use the Gram-Schmidt Orthogonalization Process for these basic solutions to obtain an orthonormal basis for the eigenspace of $\mathbf{A}$ corresonding to λ. In this process, the entries of the basis vectors remain real.

We replace the n columns of the unitary matrix $\mathbf{U}$ by n eigenvectors of $\mathbf{A}$ which form an orthonormal set in $\mathbb{R}^{n\times n}$. Let $\mathbf{C} \in \mathbb{R}^{n\times n}$ denote the matrix with these columns arranged in the same order as the columns of $\mathbf{U}$. Then $\mathbf{D} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$. □

The spectral theorem says that given a normal matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ or a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, there is a unitary matrix $\mathbf{U}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{D} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$, that is, $\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{D}$. Let us write the matrix $\mathbf{U}$ in terms of its n columns $\mathbf{U} = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{bmatrix}$ and let $\mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

Then the equation

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

means $\mathbf{A}\mathbf{u}_1 = \lambda_1\mathbf{u}_1, \ldots, \mathbf{A}\mathbf{u}_n = \lambda_n\mathbf{u}_n$, so that $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are eigenvectors of $\mathbf{A}$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$ respectively. We may list the eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ in some other order to form another unitary matrix and correspondingly change the ordering of the eigenvalues $\lambda_1, \ldots, \lambda_n$ to form another diagonal matrix which will serve the same purpose.

Since the eigenvalues $\lambda_1, \ldots, \lambda_n$ of $\mathbf{A}$ may not be distinct, we may pool together all eigenvectors among $\mathbf{u}_1, \ldots, \mathbf{u}_n$ corresponding to the same eigenvalue.

Also, since $\mathbf{A}$ is diagonalizable, the geometric multiplicity of any eigenvalue of $\mathbf{A}$ is equal to its algebraic multiplicity.

Thus if we know the eigenvalues of $\mathbf{A}$, then we may use the following procedure to form a unitary matrix $\mathbf{U}$ whose n columns are eigenvectors of $\mathbf{A}$.

1. Let $\mu_1, \ldots, \mu_k$ be the distinct eigenvalues of $\mathbf{A}$. These are the distinct roots of the characteristic polynomial of $\mathbf{A}$. Let the algebraic multiplicity of μ_j be m_j, so that $m_1 + \cdots + m_k = n$. Also,the geometric multiplicity g_j of μ_j is equal to m_j, and so $g_1 + \cdots + g_k = n$.

If in fact, $\mathbf{A}$ is self-adjoint, then each μ_j is real.

2. For each $j = 1, \ldots, k$, find a basis for the null space $\mathcal{N}(\mathbf{A} - \mu_j \mathbf{I})$ consisting of g_j elements by solving the homogeneous linear system $(\mathbf{A} - \mu_j \mathbf{I})\mathbf{x} = \mathbf{0}$ using the Gauss Elimination Method.

3. For each $j = 1, \ldots, k$, obtain an ordered orthonormal basis $(\mathbf{u}_{j,1}, \ldots, \mathbf{u}_{j,g_j})$ for $\mathcal{N}(\mathbf{A} - \mu_j \mathbf{I})$ using the Gram-Schmidt Orthonormalization Process.

4. Form an $n \times n$ matrix $\mathbf{U}$ as follows:

$$\mathbf{U} := \begin{bmatrix} \mathbf{u}_{11} & \ldots & \mathbf{u}_{1g_1} & \mathbf{u}_{21} & \ldots & \mathbf{u}_{2g_2} & \ldots & \ldots & \mathbf{u}_{k1} & \ldots \mathbf{u}_{kg_k} \end{bmatrix}$$

Now since $\mathbf{A}$ is a normal matrix, $\mathbf{u}_{ik} \perp \mathbf{u}_{j\ell}$ if $i \neq j$. Thus the n columns of $\mathbf{U}$ form an orthonormal set. Hence $\mathbf{U}$ is unitary.

Form an $n \times n$ diagonal matrix $\mathbf{D}$ as follows:

$$\mathbf{D} := \operatorname{diag}(\lambda_{11}, \ldots, \lambda_{1g_1}, \lambda_{21}, \ldots, \lambda_{2g_2}, \ldots, \ldots, \lambda_{k1}, \ldots, \lambda_{kg_k}),$$

where $\lambda_{11} = \cdots = \lambda_{1g_1} = \mu_1, \ldots, \lambda_{k1} = \cdots = \lambda_{kg_k} = \mu_k$.

Then $\boxed{\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D}}$.

A similar procedure works for a real symmetric matrix $\mathbf{A}$, and so we can find an orthogonal matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$ such that $\boxed{\mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C} = \mathbf{D}}$.

Example

Let $\mathbf{A} := \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$. Clearly, $\mathbf{A}^* = \mathbf{A}$. So $\mathbf{A}$ is self-adjoint.

1. $p_{\mathbf{A}}(t) := \det(\mathbf{A} - t\mathbf{I}) = (3-t)^2(3+t)$. Hence $\mu_1 = 3$ with $m_1 = g_1 = 2$, and $\mu_2 = -3$ with $m_2 = g_2 = 1$.

2. (i) $(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{0}$, that is,

$$\begin{bmatrix} -2 & 2 & -2 \\ 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} -2 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\iff x_1 - x_2 + x_3 = 0$ by the Gauss Elimination Method.

Hence $\mathbf{x}_{11} := \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^\mathsf{T}$ and $\mathbf{x}_{12} := \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^\mathsf{T}$ form a basis for the null space $\mathcal{N}(\mathbf{A} - 3\mathbf{I})$.

(ii) Similarly, $\mathbf{x}_{21} := \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}^\mathsf{T}$ forms a basis for the null space $\mathcal{N}(\mathbf{A} + 3\mathbf{I})$.

3. Gram-Schmidt Orthogonalization Process gives $\mathbf{u}_{11} := \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}^{\mathsf{T}}$ and $\mathbf{u}_{12} := \begin{bmatrix} -1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \end{bmatrix}^{\mathsf{T}}$, which form an orthonormal basis for $\mathcal{N}(\mathbf{A} - 3\mathbf{I})$.

Also, $\mathbf{u}_{21} := \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}^{\mathsf{T}}$ forms an orthonormal basis for $\mathcal{N}(\mathbf{A} + 3\mathbf{I})$.

4. Let $\mathbf{U} := \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$

and $\mathbf{D} := \mathrm{diag}(3, 3, -3)$. Then $\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D}$.

We now show that the unitary matrix $\mathbf{U}$ and the diagonal matrix $\mathbf{D}$ satisfying $\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D}$ we have found are not unique.

For example, let us interchange the order of the columns of **U** and make a corresponding interchange in the diagonal entries of **D**.

Thus $\mathbf{U} := \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \end{bmatrix}$

and $\mathbf{D} := \mathrm{diag}(3, -3, 3)$ would also satisfy $\mathbf{U}^*\mathbf{AU} = \mathbf{D}$.

Further, we could have chosen $\mathbf{x}_{11} := \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^\mathsf{T}$ and $\mathbf{x}_{12} := \begin{bmatrix} -1 & 1 & 2 \end{bmatrix}^\mathsf{T}$ as basis vectors for the null space $\mathcal{N}(\mathbf{A} - 3\mathbf{I})$, and orthonormalized them to obtain $\mathbf{u}_{11} := \begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^\mathsf{T}$ & $\mathbf{u}_{12} := \begin{bmatrix} -2/\sqrt{6} & -1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}^\mathsf{T}$.

Then $\mathbf{U} := \begin{bmatrix} 0 & -2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$

and $\mathbf{D} := \mathrm{diag}(3, 3, -3)$ would also satisfy $\mathbf{U}^*\mathbf{AU} = \mathbf{D}$.