Linear Algebra
Lecture 16

Saurav Bhaumik
Department of Mathematics
IIT Bombay
Spring 2025

Examples

(1) Let $\mathbf{A} \in \mathbb{C}^{3\times 3}$ be such that $\mathbf{A}^2 = 6\mathbf{A}$ and $\operatorname{trace} \mathbf{A} = 12$. Let us determine the eigenvalues of $\mathbf{A}$.

Consider the polynomial $p(t) = t^2 - 6t$. Since $p(\mathbf{A}) = \mathbf{O}$, we see that $p(\lambda) = \lambda^2 - 6\lambda = \lambda(\lambda - 6) = 0$ for every eigenvalue λ of $\mathbf{A}$. Since $\operatorname{trace} \mathbf{A} = 12$, the sum of the eigenvalues of $\mathbf{A}$ (counting algebraic multiplicities) is equal to 12. Hence the eigenvalues of $\mathbf{A}$ are $6, 6, 0$, that is, 6 is an eigenvalue of $\mathbf{A}$ of algebraic multiplicity 2, and 0 is an eigenvalue of $\mathbf{A}$ of algebraic multiplicity 1.

(ii) Let $\mathbf{A} := \begin{bmatrix} \mathbf{e}_2 & \mathbf{e}_3 & \cdots & \mathbf{e}_n & \mathbf{e}_1 \end{bmatrix}$, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the basic column vectors in $\mathbb{C}^{n\times 1}$.
Then $\mathbf{A}\mathbf{e}_1 = \mathbf{e}_2, \ldots, \mathbf{A}\mathbf{e}_{n-1} = \mathbf{e}_n$ and $\mathbf{A}\mathbf{e}_n = \mathbf{e}_1$. Hence
$$\mathbf{A} \begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} & x_n \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} x_n & x_1 & x_2 & \cdots & x_{n-1} \end{bmatrix}^{\mathsf{T}}$$
for $\begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{C}^{n\times 1}$.

The matrix $\mathbf{A}$ represents a **cyclic shift to the right**. Observe that $\mathbf{A}^n = \mathbf{I}$ and also that $\mathbf{A}$ is a unitary matrix given by

$$\mathbf{A} := \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Let λ be an eigenvalue of $\mathbf{A}$. Since $\mathbf{A}^n = \mathbf{I}$, we see that $\lambda^n = 1$. Let $\omega := e^{2\pi i/n}$. Then the nth roots of 1 are $1, \omega, \omega^2, \ldots, \omega^{n-1}$. Thus $\lambda \in \{1, \omega, \omega^2, \ldots, \omega^{n-1}\}$.
Conversely, we show that each $1, \omega, \omega^2, \ldots, \omega^{n-1}$ is an eigenvalue of $\mathbf{A}$ by finding a corresponding eigenvector.

Let $\mathbf{x} := \begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{C}^{n\times 1}$. Then $\mathbf{Ax} = \lambda\,\mathbf{x}$ means $x_1 = \lambda x_2, x_2 = \lambda x_3, \dots, x_{n-1} = \lambda x_n$ and $x_n = \lambda x_1$.

For $j = 0, \dots, n-1$, define

$$\mathbf{x}_j := \begin{bmatrix} 1 & 1/\omega^j & 1/\omega^{2j} & \cdots & 1/\omega^{(n-1)j} \end{bmatrix}^{\mathsf{T}}.$$

Now $1/\omega^{(n-1)j} = \omega^j$ since $\omega^n = 1$, and so

$$\mathbf{x}_j := \begin{bmatrix} 1 & 1/\omega^j & 1/\omega^{2j} & \cdots & 1/\omega^{(n-2)j} & \omega^j \end{bmatrix}^{\mathsf{T}}.$$

Hence $\mathbf{Ax}_j = \begin{bmatrix} \omega^j & 1 & 1/\omega^j & 1/\omega^{2j} & \cdots & 1/\omega^{(n-2)j} \end{bmatrix}^{\mathsf{T}} = \omega^j \mathbf{x}_j$. Thus $\mathbf{x}_j$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue ω^j for $j = 0, 1, \dots, n-1$.

Remark: $\mathbf{A}$ is an example of a circulant matrix. In general a circulant matrix is a square matrix each of whose rows are obtained by a cyclic shift to the right of the preceding rows.

Unitarily Diagonalizable Matrix

In the last lecture we saw that if the scalars are complex numbers, then every matrix can be unitarily 'upper triangularized'. Now we take up the question: 'Which matrices can be unitarily diagonalized?'. We saw that 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 $\mathbf{A}$. Let us investigate this condition further.

Suppose $\mathbf{A} \in \mathbb{K}^{n \times n}$ is unitarily diagonalizable, and let eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of $\mathbf{A}$ form an orthonormal basis for $\mathbb{K}^{n \times 1}$. Let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues of $\mathbf{A}$, so that $\mathbf{A}\mathbf{u}_j = \lambda_j \mathbf{u}_j$ for $j = 1, \ldots, n$. Let $\mathbf{x} \in \mathbb{K}^{n \times 1}$. Then

$$\mathbf{x} = \sum_{j=1}^{n} \langle \mathbf{u}_j, \mathbf{x} \rangle \mathbf{u}_j \quad \text{and} \quad \mathbf{A}\mathbf{x} = \sum_{j=1}^{n} \langle \mathbf{u}_j, \mathbf{x} \rangle \mathbf{A}\mathbf{u}_j = \sum_{j=1}^{n} \lambda_j \langle \mathbf{u}_j, \mathbf{x} \rangle \mathbf{u}_j.$$

Thus for any $\mathbf{x} \in \mathbb{K}^{n\times 1}$, the column vector $\mathbf{A}\,\mathbf{x}$ is completely determined by the inner products $\langle \mathbf{u}_1, \mathbf{x}\rangle, \ldots, \langle \mathbf{u}_n, \mathbf{x}\rangle$, and by the eigenvalues and eigenvectors of $\mathbf{A}$. The above representation of $\mathbf{A}\,\mathbf{x}$ can be used for various purposes.
For this reason, we would like to find necessary and/or sufficient conditions under which a square matrix can be unitarily diagonalized. We introduce a new class of matrices.

Definition

A matrix $\mathbf{A} \in \mathbb{K}^{n\times n}$ is called **normal** *if it commutes with its adjoint, that is, $\mathbf{A}^*\mathbf{A} = \mathbf{A}\,\mathbf{A}^*$.*

Examples (i) If $\mathbf{A}$ is self-adjoint (i.e., $\mathbf{A}^* = \mathbf{A}$), or skew self-adjoint (i.e., $\mathbf{A}^* = -\mathbf{A}$), or unitary, then $\mathbf{A}$ is normal.
(ii) The matrix $\mathbf{A} := \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \in \mathbb{K}^{2\times 2}$ is normal, but it is neither self-adjoint, nor skew self-adjoint, nor unitary. But not every matrix is normal, as the example $\mathbf{A} := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ shows.

Proposition

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$. Then

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

Proof. For $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n \times 1}$, $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = (\mathbf{A}\mathbf{x})^*\mathbf{A}\mathbf{y} = \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{y}$ and $\langle \mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle = (\mathbf{A}^*\mathbf{x})^*\mathbf{A}^*\mathbf{y} = \mathbf{x}^*\mathbf{A}\,\mathbf{A}^*\mathbf{y}$. This shows that if $\mathbf{A}$ is normal, then $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle$ 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 $j, k = 1, \ldots, n$. Hence $\mathbf{A}^*\mathbf{A} = \mathbf{A}\,\mathbf{A}^*$, 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, $\mathbf{x} \perp \mathbf{y}$. $\square$

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 $k = 1, \ldots, n$,

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

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. $\square$

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) Suppose $\mathbf{A} \in \mathbb{K}^{n \times n}$ is unitarily diagonalizable. Then $\mathbf{D} = \mathbf{U}^*\mathbf{A}\mathbf{U}$ for some unitary matrix $\mathbf{U}$ and diagonal matrix $\mathbf{D}$. Hence $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$, and so $\mathbf{A}^* = \mathbf{U}\mathbf{D}^*\mathbf{U}^*$. Now

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

whereas

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

Since $\mathbf{D}$ is diagonal, $\mathbf{D}^*\mathbf{D} = \mathbf{D}\mathbf{D}^*$. Consequently, $\mathbf{A}$ 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 easily see that $\mathbf{B}^*\mathbf{B} = \mathbf{B}\,\mathbf{B}^*$, that is, $\mathbf{B}$ is normal. Now, since $\mathbf{B}$ is upper triangular and normal, $\mathbf{B}$ is diagonal by the previous lemma. This proves that $\mathbf{A}$ is unitarily diagonalizable. $\square$

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.

Self-adjoint matrices

Recall that $\mathbf{A} \in \mathbb{K}^{n \times n}$ is called **self-adjoint** (or **Hermitian**) if $\mathbf{A}^* = \mathbf{A}$. We shall now prove a "spectral theorem" for self-adjoint matrices. For this purpose, the following lemma is useful.

Lemma

A matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ is diagonal with all diagonal entries real if and only if it is upper triangular and self-adjoint.

Proof. Let $\mathbf{A} := \text{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Then clearly $\mathbf{A}$ is upper triangular. Also, it is self-adjoint since $\mathbf{A}^* = \text{diag}(\overline{\lambda}_1, \ldots, \overline{\lambda}_n) = \text{diag}(\lambda_1, \ldots, \lambda_n) = \mathbf{A}$.

Conversely, suppose $\mathbf{A}$ is upper triangular and self-adjoint. Then $\mathbf{A}^*$ is lower triangular. Since $\mathbf{A}^* = \mathbf{A}$, we see that $\mathbf{A}$ is in fact diagonal with all diagonal entries real. $\square$

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.

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 $\mathbf{B}$. $\square$

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. $\square$

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

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

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

Since the eigenvalues $\lambda_1, \dots, \lambda_n$ of $\mathbf{A}$ may not be distinct, we may pool together all eigenvectors among $\mathbf{u}_1, \dots, \mathbf{u}_n$ 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:

$$\mathbf{U} := \left[\mathbf{u}_{11} \quad \ldots \quad \mathbf{u}_{1g_1} \quad \mathbf{u}_{21} \quad \ldots \quad \mathbf{u}_{2g_2} \quad \ldots \quad \ldots \quad \mathbf{u}_{k1} \quad \ldots \mathbf{u}_{kg_k}\right]$$

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

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 $\boxed{\mathbf{C}^\mathsf{T}\mathbf{AC} = \mathbf{D}}$.