6.2 Diagonalization and diagonalizability for square matrices.

0. *Assumed background.*

   - What has been covered in Topics 1-5.
   - 6.1 *Eigenvalues and eigenvectors for square matrices.*

   Abstract. We introduce:—

   - the notions of diagonalization and diagonalizability for square matrices,
   - basic 'algebraic' properties of diagonalizable square matrices,
   - some necessary and sufficient conditions for diagonalization/diagonalizability.

1. **Definition. (Diagonalization of square matrices with respect to invertible matrices.)**

   Let A, U be square matrices of the same size. Suppose U is an invertible.

   Suppose $U^{-1}AU$ is a diagonal matrix. Then we say $U^{-1}AU$ is the **diagonalization of A with respect to the invertible matrix U.**
   Remark on terminology.

   (a) If D stands for the diagonal matrix, then we refer to each of the equalities '$U^{-1}AU = D$', '$A = UDU^{-1}$', '$AU = UD$', (which are themselves equvalent to each other) as a **presentation of the diagonalization of A with respect to the invertible matrix U.** We also abuse terminologies by referring to these equalities as 'diagonalization of A with respect to U'.

   (b) When we do not want to emphasize the role of U, we may refer to $U^{-1}AU$ as a 'diagonization of A', omitting the reference to U.

2. **Definition. (Diagonalizability for square matrices.)**

   Suppose A is a square matrices.

   Then we say A is **diagonalizable** *if and only if there is some diagonalization of A with respect to some invertible matrix of the same size as A.*

3. It is noted that if a square matrix is diagonalizable then it has many distinct diagonalizations. This is a consequence of the result below, and the fact that there are many permutation matrices.

   Lemma (1).

   Let A be an $(n \times n)$-square matrix. Suppose A is diagonalizable, with the diagonalization $U^{-1}AU$ with respect to some invertible $(n \times n)$-square matrix U. Then, for any $(n \times n)$-square matrix Q, if Q is a permutation matrix, then $(UQ)^{-1}A(UQ)$ is also a diagonalization of A.

4. We shall see soon that the notions of diagonalizabibility and diagonalization are closedly tied with that of eigenvalues and eigenvectors.

   As preparation, we state and prove a simple result on matrix algebra.

   Lemma (2).

   Let A be an $(n \times n)$-square matrix, $\mu_1, \mu_2, \cdots, \mu_p$ be numbers, and $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ be non-zero column vectors with n entries.
   Write $D = \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_p)$.

   Define the $(n \times p)$-matrix V by $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_p\ ]$.

   Then the statements below are logically equivalent:—

   (1) *For each $j = 1, 2, \cdots, p$, the column vector $\mathbf{v}_j$ is an eigenvector of A with eigenvalue μ_j.*

   (2) *The equality $AV = VD$ holds.*

5. **Proof of Lemma (2).**

   Let A be an $(n \times n)$-square matrix, $\mu_1, \mu_2, \cdots, \mu_p$ be numbers, and $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ be non-zero column vectors with n entries.
   Write $D = \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_p)$.

   Define the $(n \times p)$-matrix V by $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_p\ ]$.

(a) Suppose that (1) holds: *for each $j = 1, 2, \cdots, p$, the column vector $\mathbf{v}_j$ is an eigenvector of A with eigenvalue μ_j.*

Then

$$\begin{aligned} AV = [\ A\mathbf{v}_1 \mid A\mathbf{v}_2 \mid \cdots \mid A\mathbf{v}_p\ ] &= [\ \mu_1\mathbf{v}_1 \mid \mu_2\mathbf{v}_2 \mid \cdots \mid \mu_p\mathbf{v}_p\ ] \\ &= \left[\ \mu_1 V\mathbf{e}_1^{(p)} \mid \mu_2 V\mathbf{e}_2^{(p)} \mid \cdots \mid \mu_p V\mathbf{e}_p^{(p)}\ \right] \\ &= \left[\ V(\mu_1\mathbf{e}_1^{(p)}) \mid V(\mu_2\mathbf{e}_2^{(p)}) \mid \cdots \mid V(\mu_p\mathbf{e}_p^{(p)})\ \right] \\ &= V\left[\ \mu_1\mathbf{e}_1^{(p)} \mid \mu_2\mathbf{e}_2^{(p)} \mid \cdots \mid \mu_p\mathbf{e}_p^{(p)}\ \right] \\ &= V\,\mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_p) = VD \end{aligned}$$

Hence (2) holds.

(b) Suppose (2) holds: *the equality $AV = VD$ holds.*

Note that $AV = [\ A\mathbf{v}_1 \mid A\mathbf{v}_2 \mid \cdots \mid A\mathbf{v}_p\ ]$.

Also note that

$$\begin{aligned} VD &= V\,\mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_p) = V\left[\ \mu_1\mathbf{e}_1^{(p)} \mid \mu_2\mathbf{e}_2^{(p)} \mid \cdots \mid \mu_p\mathbf{e}_p^{(p)}\ \right] \\ &= \left[\ V(\mu_1\mathbf{e}_1^{(p)}) \mid V(\mu_2\mathbf{e}_2^{(p)}) \mid \cdots \mid V(\mu_p\mathbf{e}_p^{(p)})\ \right] \\ &= \left[\ \mu_1 V\mathbf{e}_1^{(p)} \mid \mu_2 V\mathbf{e}_2^{(p)} \mid \cdots \mid \mu_p V\mathbf{e}_p^{(p)}\ \right] \\ &= [\ \mu_1\mathbf{v}_1 \mid \mu_2\mathbf{v}_2 \mid \cdots \mid \mu_p\mathbf{v}_p\ ] \end{aligned}$$

Then, by the definition of matrix equality, for each $j = 1, 2, \cdots, p$, the equality $A\mathbf{v}_j = \mu_j\mathbf{v}_j$ holds. Hence $\mathbf{v}_j$ is an eigenvector of A with eigenvalue μ_j.

Therefore (1) holds.

6. **Theorem (3). (Necessary and sufficient conditions for a presentation of diagonalization.)**

Let A be an $(n \times n)$-square matrix, and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be column vectors with n entries.

Define $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$.

Suppose U is invertible. Then the statements below are logically equivalent:—

(1) *For each $j = 1, 2, \cdots, n$, the column vector $\mathbf{u}_j$ is an eigenvector of A with eigenvalue λ_j.*

(2) *$U^{-1}AU = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ is a diagonalization of A.*

7. **Proof of Theorem (3).**

Let A be an $(n \times n)$-square matrix, and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be column vectors with n entries.

Define $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$.

Suppose U is invertible.

(a) Suppose (1) holds: *for each $j = 1, 2, \cdots, n$, the column vector $\mathbf{u}_j$ is an eigenvector of A with eigenvalue λ_j.*

By Lemma (2), $AU = U\,\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

By assumption, U is invertible. Then $U^{-1}AU = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

Therefore (2) holds.

(b) Suppose (2) holds: *$U^{-1}AU = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ is a diagonalization of A.*

Then $AU = U\,\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

Now, by Lemma (2), for each $j = 1, 2, \cdots, n$, the column vector $\mathbf{u}_j$ is an eigenvector of A with eigenvalue λ_j.

8. **Example (1). (Illustrations on diagonalizations.)**

(a) Let $A = \begin{bmatrix} 13 & 30 \\ -6 & -14 \end{bmatrix}$, and $\mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, and $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2\ ]$.

i. $\mathbf{u}_1, \mathbf{u}_2$ are eigenvectors of A with respective eigenvalues $1, -2$.

ii. $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent. Then U is invertible.

Its matrix inverse is given by $U^{-1} = \begin{bmatrix} 1 & 2 \\ -2 & -5 \end{bmatrix}$.

iii. By direct verification, we see that $U^{-1}AU = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$.

(b) Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, and $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$, and $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3\ ]$.

i. $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are eigenvectors of A with respective eigenvalues $1, 2, 3$.

ii. $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly independent. Then U is invertible.

Its matrix inverse is given by $U^{-1} = \begin{bmatrix} 1 & -1 & 1/2 \\ 0 & 1 & -2 \\ 0 & 0 & 1/2 \end{bmatrix}$.

iii. By direct verification, we see that $U^{-1}AU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

(c) Let $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & 1 & 1 & 1 \\ -5 & 2 & 5 & -1 \end{bmatrix}$, and $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 5 \\ -1 \\ -5 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} 1 \\ -5 \\ -3 \\ 15 \end{bmatrix}$, and $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3 \mid \mathbf{u}_4\ ]$.

i. $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ are eigenvectors of A with respective eigenvalues $1, -1, 3, -3$.

ii. $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ are linearly independent. Then U is invertible.

Its matrix inverse is given by $U^{-1} = \begin{bmatrix} 5/8 & -1/4 & 0 & -1/8 \\ 1/4 & 1/8 & -1/8 & 0 \\ 0 & 1/8 & 5/24 & 1/12 \\ 1/8 & 0 & -1/12 & 1/24 \end{bmatrix}$.

iii. By direct verification, we see that $U^{-1}AU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$.

(d) Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$, and $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, and $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3\ ]$.

i. $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are eigenvectors of A with respective eigenvalues $4, 1, 1$.

ii. $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly independent. Then U is invertible.

Its matrix inverse is given by $U^{-1} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & -2/3 & 1/3 \\ 1/3 & 1/3 & -2/3 \end{bmatrix}$.

iii. By direct verification, we see that $U^{-1}AU = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

9. Using the 'dictionary' about equivalent formulations of invertibility, or specifically, the logical equivalence between the invertibility of a square matrix and linear independence of the columns of the square matrix concerned, we obtain the result below. (This reason has already been hinted by Example (1).)

Theorem (4). (Necessary and sufficient conditions for diagonalizability.)

Suppose A is an $(n \times n)$-square matrix.

Then the statements below are logically equivalent:—

(1) *There are n linearly independent eigenvectors of A.*

(2) *A is diagonalizable.*

Now suppose either of (1), (2) *holds. (So both of them hold.)*

Further suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are n linearly independent eigenvectors of A, with respective eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$.

Define $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$.

Then $U^{-1}AU = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ is a diagonalization of A.

Remark. Theorem (3) and Theorem (4) are almost identical in content. The only difference is about where we put the emphasis. In Theorem (3), we focus on the question whether something is a diagonalization for a given matrix. In Theorem (4), we focus on the question whether a given matrix admits any diagonalization.

The proof of Theorem (4) is very similar to that of Theorem (3), and hence is left as an exercise.

10. Using Theorem (3) and Theorem (4), we can deduce the result below:—

Theorem (5).

Let A be an $(n \times n)$-square matrix. Suppose A is diagonalizable, with a diagonalization $U^{-1}AU = D$, for some invertible $(n \times n)$-square matrix U and for some $(n \times n)$-diagonal matrix D.

Then the statements below hold:—

(1) *For each positive integer p, the matrix A^p is diagonalizable, with a diagonalization given by $U^{-1}A^pU = D^p$.*

(2) *Suppose A is invertible. Then D is invertible, and every diagonal entry of D is non-zero.*
Moreover, A^{-1} is diagonalizable, with a diagonalization given by $U^{-1}A^{-1}U = D^{-1}$.

(3) *A^t is diagonalizable, with a diagonalization given by $U^tA^t(U^{-1})^t = D$.*
Moreover, for each $j = 1, 2, \cdots, n$, the j-th column of $(U^{-1})^t$ is an eigenvector of A with its corresponding eigenvalue being the j-th diagonal entry of D.

Proof of Theorem (5). Exercise (in matrix algebra on product and powers, inverse and transpose).

Remark. For each $j = 1, 2, \cdots, n$, denote the j-th column of U by $\mathbf{u}_j$, and the j-th diagonal entry of D by λ_j.

(a) Immediate from the statements (1), (2) in the conclusion are:—

i. For each positive integer p, $\mathbf{u}_j$ is an eigenvector of A^p with eigenvalue ${\lambda_j}^p$.
It will further follow that if $\mathbf{v}$ is an eigenvector of A then $\mathbf{v}$ is an eigenvector of A^p for each positive integer p.

ii. $\mathbf{u}_j$ is an eigenvector of A^{-1} with eigenvalue $1/\lambda_j$.
It will further follow that if $\mathbf{v}$ is an eigenvector of A then $\mathbf{v}$ is an eigenvector of A^{-1}.

(b) Immediate from the statement (3) is that A and A^t have the same collection of eigenvalues: $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of A^t.
Be aware that it does not follow that $\mathbf{u}_j$ is an eigenvector of A^t with eigenvalue λ_j.
All that can be said is that the j-th column of $(U^{-1})^t$ is an eigenvector of A^t with eigenvalue λ_j.

11. **Example (2). (Application of Theorem (5) in the computation of positive powers of diagonalizable square matrices.)**

(a) Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$.

A is diagonalizable, with a diagonalization given by $U^{-1}AU = D$, in which $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3\ ]$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$, and $D = \mathrm{diag}(1, 2, 3)$.

i. Note that $A = UDU^{-1}$, and $U = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{bmatrix}$ and $U^{-1} = \begin{bmatrix} 1 & -1 & 1/2 \\ 0 & 1 & -2 \\ 0 & 0 & 1/2 \end{bmatrix}$.

ii. For each positive integer p, we have

$$\begin{aligned} A^p &= UD^pU^{-1} = U(\mathrm{diag}(1, 2^p, 3^p))U^{-1} \\ &= \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^p & 0 \\ 0 & 0 & 3^p \end{bmatrix} \begin{bmatrix} 1 & -1 & 1/2 \\ 0 & 1 & -2 \\ 0 & 0 & 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 + 2^p & 1/2 - 2 \cdot 2^p + (3/2) \cdot 3^p \\ 0 & 2^p & -2 \cdot 2^p + 2 \cdot 3^p \\ 0 & 0 & 3^p \end{bmatrix}. \end{aligned}$$

(b) Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

A is diagonalizable, with a diagonalization given by $U^{-1}AU = D$, in which $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3\ ]$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, and $D = \mathrm{diag}(4, 1, 1)$.

i. Note that $A = U\,\mathrm{diag}(4,1,1)U^{-1}$, and $U = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$ and $U^{-1} = \dfrac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$.

ii. For each positive integer p, we have

$$\begin{aligned} A^p &= UD^pU^{-1} = U(\mathrm{diag}(4^p,1,1))U^{-1} \\ &= \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} 4^p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 4^p+2 & 4^p-1 & 4^p-1 \\ 4^p-1 & 4^p+2 & 4^p-1 \\ 4^p-1 & 4^p-1 & 4^p+2 \end{bmatrix}. \end{aligned}$$

12. With an argument similar to that for Theorem (5), we can also deduce the result below:—

Theorem (6).

Let A, B be $(n \times n)$-square matrices. Suppose A, B are diagonalizable, with diagonalizations with respect to a common invertible $(n \times n)$-square matrix say, U.

Then the statements below hold:—

(1) *There are some $(n \times n)$-diagonal matrices D, E so that $U^{-1}AU = D$ and $U^{-1}BU = E$.*

(2) *For any numbers α, β, the square matrix $\alpha A + \beta B$ is diagonalizable, with a diagonalization given by $U^{-1}(\alpha A + \beta B)U = D + E$.*

(3) *AB diagonalizable, with a diagonalization given by $U^{-1}ABU = DE$.*

13. As suggested by examples that we have seen on eigenvalues and eigenvectors for upper-triangular matrices, we know that there is no guarantee for an arbitrary square matrix to have sufficiently many linearly independent eigenvectors for it to be diagonalizable.

Example (3). (Non-diagonalizable square matrices.)

(a) Let $A = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$.

The only eigenvalue of A is 1.

The only eigenvectors of A are $\mathbf{u}_1 = \mathbf{e}_1^{(2)}$, corresponding to this eigenvalue, and its non-zero scalar multiples. Then A is not diagonalizable.

(b) Let $A = \begin{bmatrix} 1 & 4 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$.

The only eigenvalue of A is 1.

The only eigenvectors of A are $\mathbf{u}_1 = \mathbf{e}_1^{(3)}$, corresponding to this eigenvalue, and its non-zero scalar multiples. Then A is not diagonalizable.

14. The 'relation' between a diagonalizable square matrix and the diagonal matrix involved in a presentation of diagonalization is an instance of something more general, known as similarity for square matrices.

Definition. (Similarity for square matrices.)

Let A, B be $(n \times n)$-square matrices.

We say that A is **similar** *to B if and only if the statement* (SI) *holds:*

(SI) *there exists some invertible $(n \times n)$-square matrix U such that $A = U^{-1}BU$.*

Theorem (7). (Similarity as an 'equivalence relation'.)

The statements below hold:—

(1) *Suppose A is an $(n \times n)$-square matrix. Then A is similar to A.*

(2) *Let A, B be $(n \times n)$-square matrices. Suppose A is similar to B. Then B is similar to A.*

(3) *Let A, B, C be $(n \times n)$-square matrices. Suppose A is similar to B, and B is similar to C. Then A is similar to C.*

15. We have seen that some square matrices with real entries may have non-real complex numbers as eigenvalues, which in turn correspond to eigenvectors with entries involving complex numbers.

The theory developed for diagonalizations and diagonalizability applies to these matrices as well, (as long as we agree to think of real numbers as complex numbers).

Example (4). (Diagonalization of a square matrix involving the use of complex numbers.)

Let $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, and $\mathbf{u}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$, and $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2\ ]$

(a) i. $\mathbf{u}_1, \mathbf{u}_2$ are eigenvectors of A with respective eigenvalues $1+i, 1-i$.

ii. $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent (over the complex numbers). Then U is invertible.

Its matrix inverse is given by $U^{-1} = \dfrac{1}{2}\begin{bmatrix} -i & 1 \\ i & 1 \end{bmatrix}$

iii. $U^{-1}AU = \begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix}$, and $A = U\begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix}U^{-1}$.

(b) For each positive integer p, we have

$$\begin{aligned} A^p &= U\begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix}^p U^{-1} \\ &= \frac{1}{2}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}\begin{bmatrix} (1+i)^p & 0 \\ 0 & (1-i)^p \end{bmatrix}\begin{bmatrix} -i & 1 \\ i & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} (1+i)^p + (1-i)^p & i(1+i)^p - i(1-i)^p \\ -i(1+i)^p + i(1-i)^p & (1+i)^p + (1-i)^p \end{bmatrix} \end{aligned}$$

Although the respective presentations of the entries of A^p involve the use of complex numbers, the entries themselves are in fact all real. (This is expected because the entries of A are all real.)