2. EIGENVALUES, EIGENVECTORS AND DIAGONALIZABILITY

Definition 36 *An $n \times n$ matrix A is said to be **diagonalizable** if there is an invertible matrix P such that $P^{-1}AP$ is diagonal.*

Two questions immediately spring to mind: why might this be a useful definition, and how might we decide whether such a matrix P exists? In an attempt to partially answer the first question, we note

$$(P^{-1}AP)^k = (P^{-1}AP)(P^{-1}AP) \times \cdots \times (P^{-1}AP) = P^{-1}A^kP,$$

as all the internal products PP^{-1} cancel. Thus if $P^{-1}AP = D$ is diagonal then

$$A^k = PD^kP^{-1} \qquad \text{for a natural number } k$$

and so we are in a position to easily calculate the powers of A. So ease of calculation is clearly one advantage of a matrix being diagonalizable.

For now we will consider this reason enough to seek to answer the second question: how do we determine whether such a P exists? Suppose such a P exists and has columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. As P is invertible then the $\mathbf{v}_i$ are independent. Further as $AP = PD$ where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ then we have that

$$i\text{th column of } AP = A\mathbf{v}_i \qquad \text{and} \qquad i\text{th column of } PD = P(\lambda_i \mathbf{e}_i^T) = \lambda_i \mathbf{v}_i.$$

So the columns of P are n independent vectors, each of which A maps to a scalar multiple of itself. Thus we make the following definitions.

Definition 37 *Let A be an $n \times n$ matrix. We say that a non-zero vector $\mathbf{v}$ in $\mathbb{R}^n_{\text{col}}$ is an **eigenvector** [1] of A if $A\mathbf{v} = \lambda\mathbf{v}$ for some scalar λ. The scalar λ is called the **eigenvalue** of $\mathbf{v}$ and we will also refer to $\mathbf{v}$ as a λ-**eigenvector**.*

Definition 38 *n linearly independent eigenvectors of an $n \times n$ matrix A are called an **eigenbasis**.*

Remark 39 *As the determinant of a linear map $T\colon V \to V$ of a finite-dimensional vector space is well-defined, then we can make the same definitions of eigenvalue, eigenvector and eigenbasis for T.*

And we have partly demonstrated the following.

[1] The German adjective *eigen* means 'own' or 'particular'. David Hilbert was the first to use the term in the early 20th century. The term *proper* or *characteristic* is sometimes also used, especially in older texts.

Theorem 40 *An $n \times n$ matrix A is diagonalizable if and only if A has an eigenbasis.*

Proof. We showed above that if such a P exists then its columns form an eigenbasis. Conversely if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form an eigenbasis, with respective eigenvalues $\lambda_1, \ldots, \lambda_n$, we define

$$P = (\mathbf{v}_1 | \cdots | \mathbf{v}_n)$$

to be the $n \times n$ matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Again P is invertible as its columns are linearly independent. Then

$$P\mathbf{e}_i^T = \mathbf{v}_i \qquad \text{and} \qquad AP\mathbf{e}_i^T = A\mathbf{v}_i = \lambda_i \mathbf{v}_i = \lambda_i P\mathbf{e}_i^T = P(\lambda_i \mathbf{e}_i^T),$$

so that $P^{-1}AP\mathbf{e}_i^T = \lambda_i \mathbf{e}_i^T$ for each i or equivalently

$$P^{-1}AP = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

■

Note that λ is an eigenvalue of A if and only if the equation $A\mathbf{v} = \lambda\mathbf{v}$ has a non-zero solution or equivalently if $(\lambda I_n - A)\mathbf{v} = \mathbf{0}$ has a non-zero solution. This is equivalent to $\lambda I_n - A$ being singular, which in turn is equivalent to $\det(\lambda I_n - A) = 0$. Thus we have shown (a) below.

Proposition 41 *Let A be an $n \times n$ matrix.*
(a) A real number λ is an eigenvalue of A if and only if $x = \lambda$ is a root of $\det(xI_n - A) = 0$.
(b) $\det(xI_n - A)$ is a polynomial in x of degree n which is monic (i.e. leading coefficient is 1).
(c) If $\det(xI_n - A) = x^n + c_{n-1}x^{n-1} + \cdots + c_0$ then

$$c_0 = (-1)^n \det A \qquad \textit{and} \qquad c_{n-1} = -\operatorname{trace}(A).$$

Proof. (b) Note

$$\det(xI_n - A) = \begin{vmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x - a_{nn} \end{vmatrix}.$$

This determinant is the sum of $n!$ products that take one entry from each row and each column. The largest power of x is produced from the product of the diagonal entries

$$(x - a_{11})(x - a_{22}) \cdots (x - a_{nn}). \tag{2.1}$$

The greatest power of x here is x^n and the coefficient of x^n is 1. All other products give polynomials in x of degree strictly less than n.

(c) By setting $x = 0$ we see that

$$c_0 = \det(-A) = (-1)^n \det A.$$

Contributions to the x^{n-1} term only come from the product of the diagonal entries (2.1). If one diagonal entry is omitted from a product then necessarily a second diagonal entry is also omitted and thus the greatest power of x from such a product can be x^{n-2}. The coefficient of x^{n-1} from (2.1) is

$$-a_{11} - a_{22} - \cdots - a_{nn} = -\operatorname{trace}(A).$$

■

Definition 42 *Let A be a real $n \times n$ matrix. Then the* ***characteristic polynomial*** *of A is*

$$\chi_A(x) = \det(xI_n - A).$$

Example 43 *Find the eigenvalues of the following matrices.*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}; \qquad B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}; \qquad C = \begin{pmatrix} 3 & 2 & -4 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{pmatrix}; \qquad D = \begin{pmatrix} 5 & -3 & -5 \\ 2 & 9 & 4 \\ -1 & 0 & 7 \end{pmatrix}.$$

Solution. By Proposition 41(a) this is equivalent to finding the *real* roots of the matrices' characteristic polynomials.

(a) The eigenvalues of A are 0 and 2 as

$$\chi_A(x) = \begin{vmatrix} x-1 & -1 \\ -1 & x-1 \end{vmatrix} = (x-1)^2 - 1 = x(x-2).$$

(b) Similarly note

$$\chi_B(x) = \begin{vmatrix} x-1 & 1 \\ -1 & x-1 \end{vmatrix} = (x-1)^2 + 1 = x^2 - 2x + 2.$$

Now $\chi_B(x)$ has no real roots (the roots are $1 \pm i$) and so B has no eigenvalues.

(c) As C is triangular then we can immediately see that $\chi_C(x) = (x-3)(x-1)(x-3)$. So C has eigenvalues $1, 3, 3$, the eigenvalue of 3 being a repeated root of $\chi_C(x)$.

(d) Finally D has eigenvalues $6, 6, 9$, the eigenvalue of 6 being repeated as $\chi_D(x)$ equals

$$\begin{vmatrix} x-5 & 3 & 5 \\ -2 & x-9 & -4 \\ 1 & 0 & x-7 \end{vmatrix} = \begin{vmatrix} x-6 & x-6 & x-6 \\ -2 & x-9 & -4 \\ 1 & 0 & x-7 \end{vmatrix} = (x-6) \begin{vmatrix} 1 & 1 & 1 \\ -2 & x-9 & -4 \\ 1 & 0 & x-7 \end{vmatrix}$$

$$= (x-6) \begin{vmatrix} 1 & 0 & 0 \\ -2 & x-7 & -2 \\ 1 & -1 & x-8 \end{vmatrix} = (x-6) \begin{vmatrix} x-7 & -2 \\ -1 & x-8 \end{vmatrix} = (x-6)^2(x-9).$$

■

Here follow some basic facts about eigenvalues, eigenvectors and diagonalizability.

Proposition 44 *Let A be an $n \times n$ matrix and $\lambda \in \mathbb{R}$.*
(a) The λ-eigenvectors of A, together with $\mathbf{0}$, form a subspace of $\mathbb{R}^n_{\text{col}}$. This is called the λ- ***eigenspace****.*
(b) For $1 \leqslant i \leqslant k$, let $\mathbf{v}_i$ be a λ_i-eigenvector of A. If $\lambda_1, \ldots, \lambda_k$ are distinct then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are independent.

Proof. (a) This is $\ker(A - \lambda I_n)$ and kernels are subspaces.

(b) may be proven by induction as follows. Note that $\mathbf{v}_1 \neq \mathbf{0}$ (as it is an eigenvector) and so $\mathbf{v}_1$ makes an independent set. Suppose, as our inductive hypothesis, that $\mathbf{v}_1, \ldots, \mathbf{v}_i$ are linearly independent vectors and that

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_i \mathbf{v}_i + \alpha_{i+1} \mathbf{v}_{i+1} = \mathbf{0} \tag{2.2}$$

for some reals $\alpha_1, \ldots, \alpha_{i+1}$. If we apply A to both sides of (2.2), we find

$$\alpha_1\lambda_1\mathbf{v}_1 + \cdots + \alpha_i\lambda_i\mathbf{v}_i + \alpha_{i+1}\lambda_{i+1}\mathbf{v}_{i+1} = \mathbf{0}. \tag{2.3}$$

Now subtracting λ_{i+1} times (2.2) from (2.3) we arrive at

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \cdots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i = \mathbf{0}.$$

By hypothesis $\mathbf{v}_1, \ldots, \mathbf{v}_i$ are linearly independent vectors and hence $\alpha_j(\lambda_j - \lambda_{i+1}) = 0$ for $1 \leqslant j \leqslant i$. As $\lambda_1, \ldots, \lambda_i$ are distinct then $\alpha_j = 0$ for $1 \leqslant j \leqslant i$ and then by (2.2) $\alpha_{i+1} = 0$. We have shown that $\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}$ are linearly independent vectors and so (b) follows by induction.

■

Corollary 45 *If an $n \times n$ matrix has n distinct eigenvalues then it is diagonalizable.*

Proof. Let $\lambda_1, \ldots, \lambda_n$ denote the distinct eigenvalues. For each i there is a λ_i-eigenvector $\mathbf{v}_i$ and by Proposition 44(b) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are independent. There being n of them they form an eigenbasis. ■

- It is important to note this is a sufficient, but not a necessary condition for diagonalizability. For example, I_n is diagonal (and so diagonalizable) but has eigenvalue 1 repeated n times.

Example 46 *Determine the eigenvectors and diagonalizability of the matrices A, B, C, D from Example 43.*

Solution. (a) We determined that A has eigenvalues $\lambda = 0$ and 2. Note that

$$\begin{aligned} \lambda &= 0: \qquad \ker\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle; \\ \lambda &= 2: \qquad \ker\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle. \end{aligned}$$

$(1, -1)^T$ and $(1, 1)^T$ form an eigenbasis and if we set

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \qquad \text{then} \qquad P^{-1}AP = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Note that we could have created an invertible matrix P by swapping its columns and we would have found $P^{-1}AP = \operatorname{diag}(2, 0)$. The eigenvalues appear in the diagonal of $P^{-1}AP$ in the order the corresponding eigenvectors appear in the columns of P.

(b) B has no real eigenvalues and so no eigenvectors. Consequently B is not diagonalizable. (At least not using a real matrix P; however see Example 47.)

(c) C has eigenvalues $1, 3, 3$. Note

$$\begin{aligned} \lambda &= 3: \qquad \ker\begin{pmatrix} 0 & 2 & -4 \\ 0 & -2 & 4 \\ 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\rangle. \\ \lambda &= 1: \qquad \ker\begin{pmatrix} 2 & 2 & -4 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\rangle. \end{aligned}$$

An eigenbasis is $(1,0,0)^T$, $(0,2,1)^T$ and $(1,-1,0)^T$. Setting

$$P=\begin{pmatrix}1&0&1\\0&2&-1\\0&1&0\end{pmatrix}\qquad\text{then}\qquad P^{-1}CP=\operatorname{diag}(3,3,1).$$

(d) D has eigenvalues $6,6,9$. Note that

$$\begin{aligned}\lambda&=6: &\quad \ker\begin{pmatrix}-1&-3&-5\\2&3&4\\-1&0&1\end{pmatrix}&=\left\langle\begin{pmatrix}1\\-2\\1\end{pmatrix}\right\rangle.\\ \lambda&=9: &\quad \ker\begin{pmatrix}-4&-3&-5\\2&0&4\\-1&0&-2\end{pmatrix}&=\left\langle\begin{pmatrix}-2\\1\\1\end{pmatrix}\right\rangle\end{aligned}$$

The 6-eigenvectors are non-zero multiples of $(1,-2,1)^T$ and the 9-eigenvectors are non-zero multiples of $(-2,1,1)^T$. As we can find no more than two independent eigenvectors, then there is no eigenbasis and D is not diagonalizable. In fact, we will shortly see that as soon as we noted the multiplicity two eigenvalue 6 yielded only one independent eigenvector then D could not be diagonalizable. ■

Example 47 *Find a complex matrix P such that $P^{-1}BP$ is diagonal, where B is as given in Example 43.*

Remark 48 *When we defined 'diagonalizability' in Definition 36 we were, strictly speaking, defining 'diagonalizability over $\mathbb{R}$'. We would say that B is not diagonalizable over $\mathbb{R}$ as no such matrix P with real entries exists, but B is diagonalizable over $\mathbb{C}$ as such a complex matrix P does exist.*

Solution. The roots of $\chi_B(x)=(x-1)^2+1$ are $1\pm i$. When the field of scalars is $\mathbb{C}$, then these are distinct complex eigenvalues and we know that B is diagonalizable over $\mathbb{C}$. Note that

$$\begin{aligned}\lambda&=1+i: &\quad \ker\begin{pmatrix}-i&-1\\1&-i\end{pmatrix}&=\left\langle\begin{pmatrix}i\\1\end{pmatrix}\right\rangle;\\ \lambda&=1-i: &\quad \ker\begin{pmatrix}i&-1\\1&i\end{pmatrix}&=\left\langle\begin{pmatrix}1\\i\end{pmatrix}\right\rangle.\end{aligned}$$

So we may take

$$P=\begin{pmatrix}i&1\\1&i\end{pmatrix}\qquad\text{and find}\qquad P^{-1}BP=\begin{pmatrix}1+i&0\\0&1-i\end{pmatrix}.$$

■

Examples 43 and 46 cover the various eventualities that may arise when investigating the diagonalizability of matrices. In summary, the checklist when testing a matrix for diagonalizability is as follows.

Algorithm 49 ***(Determining Diagonalizability over*** $\mathbb{R}$ ***and*** $\mathbb{C}$*)*
(a) Let A be an $n \times n$ matrix. Determine its characteristic polynomial χ_A.
(b) If any of the roots of χ_A are not real, then A is not diagonalizable over $\mathbb{R}$.
(c) If all the roots of χ_A are real and distinct then A is diagonalizable over $\mathbb{R}$.
(d) If all the roots of χ_A are real, and for each root λ there are as many independent λ-eigenvectors as repeated factors of $x - \lambda$ in $\chi_A(x)$, then A is diagonalizable over $\mathbb{R}$.
(c)' If the roots of χ_A are distinct complex numbers then A is diagonalizable over $\mathbb{C}$.
(d)' If for each root λ of χ_A and there are as many independent λ-eigenvectors in $\mathbb{C}^n_{\text{col}}$ as repeated factors of $x - \lambda$ in $\chi_A(x)$, then A is diagonalizable over $\mathbb{C}$.

(c), and the same result (c') for complex matrices, were proven in Corollary 45.
(d), and it complex version (d'), will be proven in Corollary 52.

So a square matrix can fail to be diagonalizable using a real invertible matrix P when

- not all the roots of $\chi_A(x)$ are real – counting multiplicities and including complex roots, $\chi_A(x)$ has n roots. However we will see (Proposition 51) that there are at most as many independent λ-eigenvectors as repetitions of λ as a root. So if some roots are not real we cannot hope to find n independent real eigenvectors. This particular problem can be circumvented by seeking an invertible complex matrix P instead.

- some (real or complex) root λ of $\chi_A(x)$ has fewer independent λ-eigenvectors (in $\mathbb{R}^n_{\text{col}}$ or $\mathbb{C}^n_{\text{col}}$) than there are factors of $x - \lambda$ in $\chi_A(x)$.

The latter problem cannot be circumvented, however this latter possibility is reassuringly unlikely. If a matrix's entries contain experimental data or randomly selected entries – rather than being a contrived exercise – then $\chi_A(x)$ will almost certainly have distinct complex roots and so A will be diagonalizable using a complex invertible matrix P.

Definition 50 *Let A be an $n \times n$ matrix with eigenvalue λ.*
(a) The ***algebraic multiplicity*** *of λ is the number of factors of $x - \lambda$ in the characteristic polynomial $\chi_A(x)$.*
(b) The ***geometric multiplicity*** *of λ is the maximum number of linearly independent λ-eigenvectors. This equals the dimension of the λ-eigenspace.*

Proposition 51 *The geometric multiplicity of an eigenvalue is less than or equal to its algebraic multiplicity.*

Proof. Let g and a respectively denote the geometric and algebraic multiplicities of an eigenvalue λ of an $n \times n$ matrix A. There are then g independent λ-eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_g$ which we can extend these vectors to n independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. If we put $\mathbf{v}_1, \ldots, \mathbf{v}_n$ as the columns of a matrix P then, arguing as in Theorem 40, we have

$$P^{-1}AP = \begin{pmatrix} \lambda I_g & B \\ 0 & C \end{pmatrix},$$

where B is a $g\times(n-g)$ matrix and C is $(n-g)\times(n-g)$. By the product rule for determinants we have

$$\begin{aligned}
\chi_A(x) &= \det(xI_n - A) \\
&= \det(P(xI_n - P^{-1}AP)P^{-1}) \\
&= \det(xI_n - P^{-1}AP) \\
&= \det\begin{pmatrix} (x-\lambda)I_g & -B \\ 0 & xI_{n-g}-C \end{pmatrix} \\
&= (x-\lambda)^g\chi_C(x).
\end{aligned}$$

So there are at least g factors of $x-\lambda$ in $\chi_A(x)$ and hence $a \geqslant g$. ■

Corollary 52 *Let A be a square matrix with all the roots of χ_A being real. Then A is diagonalizable if and only if, for each eigenvalue, its geometric multiplicity equals its algebraic multiplicity.*

Proof. Let the *distinct* eigenvalues of A be $\lambda_1,\ldots,\lambda_k$ with geometric multiplicities $g_1,\ldots,g_k$ and algebraic multiplicities $a_1,\ldots,a_k$. By the previous proposition

$$g_1+\cdots+g_k \leqslant a_1+\cdots+a_k = \deg\chi_A = n, \tag{2.4}$$

the equalities following as all the roots of χ_A are real. We can find g_i linearly independent λ_i-eigenvectors $\mathbf{v}_1^{(i)},\ldots,\mathbf{v}_{g_i}^{(i)}$ for each i. If $g_i=a_i$ for each i then we have n eigenvectors in all, but if $g_i<a_i$ for any i then $g_1+\cdots+g_k<n$ by (2.4), so we will not be able to find n independent eigenvectors and no eigenbasis exists. It remains to show that if $g_i=a_i$ for each i then these n eigenvectors are indeed independent. Say that

$$\sum_{i=1}^{k}\sum_{j=1}^{g_i}\alpha_j^{(i)}\mathbf{v}_j^{(i)} = \mathbf{0},$$

for some scalars $\alpha_j^{(i)}$. As $\lambda_1,\ldots,\lambda_k$ are distinct, arguing along the same lines as Proposition 44(b), it follows that

$$\mathbf{w}_i = \sum_{j=1}^{g_i}\alpha_j^{(i)}\mathbf{v}_j^{(i)} = \mathbf{0}, \qquad \text{for each } i,$$

as $\mathbf{w}_i$ is a λ_i-eigenvector (or $\mathbf{0}$). Now the vectors $\mathbf{v}_1^{(i)},\ldots,\mathbf{v}_{g_i}^{(i)}$ are independent and so $\alpha_j^{(i)}=0$ for each i and j, and hence these n vectors are indeed independent and so form an eigenbasis. ■

Remark 53 *Implicit in the above proof is the fact that the eigenspaces form a direct sum, whether or not the matrix (or linear map) is diagonalizable. From this point of view a matrix (or linear map) is diagonalizable if and only if $\mathbb{R}^n_{\text{col}}$ (or V) can be written as a direct sum of eigenspaces.*

If we recall the matrices A, B, C, D from Example 43, we can now see that A meets criterion (c) and so is diagonalizable; B meets criterion (b) and so is not diagonalizable over $\mathbb{R}$ but does meet criterion (c)' and is diagonalizable over $\mathbb{C}$; matrix C meets criterion (d) and so is diagonalizable over $\mathbb{R}$; matrix D fails criteria (c) and (d) and so is not diagonalizable over $\mathbb{R}$, specifically because the eigenvalue $\lambda = 6$ has a greater algebraic multiplicity of 2 than its geometric multiplicity of 1. This problem remains true when using complex numbers and so D is also not diagonalizable over $\mathbb{C}$.

Remark 54 *(**Diagonalizability over a general field**) We can decide on the diagonalizability of a matrix over a general field by following the same procedures as above. Firstly all the roots of the characteristic polynomial need to be in the field, and then for each eigenvalue the algebraic multiplicy needs to equal the geometric multiplicity. For example the matrix*

$$B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

has characteristic polynomial $(x-1)^2 + 1 = x^2 - 2x + 2$.

- *Over $\mathbb{C}$ this is diagonalizable as B has distinct roots $1 \pm i$.*
- *The same would be true over the field $\mathbb{Q}[i] = \{q_1 + q_2 i \mid q_1, q_2 \in \mathbb{Q}\}$.*
- *Over $\mathbb{R}$ and $\mathbb{Q}$ the characteristic polynomial has no roots and so B is not diagonalizable.*
- *Over $\mathbb{Z}_2$ the characteristic polynomial equals x^2 but the 0-eigenspace, $\left\langle (1,1)^T \right\rangle$, is 1-dimensional. As $g_0 = 1 < 2 = a_0$ then B is not diagonalizable.*
- *Over $\mathbb{Z}_3$ the characteristic polynomials has no roots as $-1 = 2$ has no square root and so B is not diagonalizable.*
- *Over $\mathbb{Z}_5$ we note $-1 = 4 = 2^2$ and so $x^2 - 2x + 2 = (x+1)(x-3)$. As B has distinct eigenvalues it is diagonalizable.*

Example 55 *Show that the matrix A below is diagonalizable and find A^n where n is a positive integer.*

$$A = \begin{pmatrix} 2 & 2 & -2 \\ 1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Solution. Adding column 2 of $xI - A$ to column 1, we can see that $\chi_A(x)$ equals

$$\begin{vmatrix} x-2 & -2 & 2 \\ -1 & x-3 & 1 \\ 1 & -1 & x-1 \end{vmatrix} = \begin{vmatrix} x-4 & -2 & 2 \\ x-4 & x-3 & 1 \\ 0 & -1 & x-1 \end{vmatrix}$$

$$= \begin{vmatrix} x-4 & -2 & 2 \\ 0 & x-1 & -1 \\ 0 & -1 & x-1 \end{vmatrix} = (x-4)(x-2)x.$$

Hence the eigenvalues are $\lambda = 0, 2, 4$. That they are distinct implies immediately that A is diagonalizable. Note

$$\begin{array}{lll}
\lambda = 0: & \ker \begin{pmatrix} -2 & -2 & 2 \\ -1 & -3 & 1 \\ 1 & -1 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle; \\
\lambda = 2: & \ker \begin{pmatrix} 0 & -2 & 2 \\ -1 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle; \\
\lambda = 4: & \ker \begin{pmatrix} 2 & -2 & 2 \\ -1 & 1 & 1 \\ 1 & -1 & 3 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle.
\end{array}$$

So three independent eigenvectors are $(1,0,1)^T$, $(0,1,1)^T$, $(1,1,0)^T$. If we set

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad \text{so that} \qquad P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix},$$

then $P^{-1}AP = \operatorname{diag}(0,2,4)$ and $P^{-1}A^nP = (P^{-1}AP)^n = \operatorname{diag}(0, 2^n, 4^n)$. Finally A^n equals

$$\begin{aligned}
& \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 4^n \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \\
= & \begin{pmatrix} 2^{2n-1} & 2^{2n-1} & -2^{2n-1} \\ 2^{2n-1} - 2^{n-1} & 2^{n-1} + 2^{2n-1} & 2^{n-1} - 2^{2n-1} \\ -2^{n-1} & 2^{n-1} & 2^{n-1} \end{pmatrix}.
\end{aligned}$$

■

Example 56 *Let*

$$A = \begin{pmatrix} 6 & 1 & 2 \\ 0 & 7 & 2 \\ 0 & -2 & 2 \end{pmatrix}.$$

(a) Show that A has two eigenvalues λ_1 and λ_2. Is A diagonalizable?
(b) Show further that $A^2 = (\lambda_1 + \lambda_2)A - \lambda_1\lambda_2 I$. Are there scalars $a_0, a_1, \ldots, a_n$, for some n, such that

$$a_n A^n + a_{n-1}A^{n-1} + \cdots + a_0 I = \operatorname{diag}(1,2,3) \text{ ?}$$

Solution. (a) We have

$$\chi_A(x) = \begin{vmatrix} x-6 & -1 & -2 \\ 0 & x-7 & -2 \\ 0 & 2 & x-2 \end{vmatrix} = (x-6)\{(x-7)(x-2)+4\} = (x-6)^2(x-3).$$

As one of the eigenvalues is repeated then we cannot immediately decide on A's diagonalizability. Investigating the repeated eigenvalue we see

$$\lambda_1 = 6: \qquad \ker \begin{pmatrix} 0 & -1 & -2 \\ 0 & -1 & -2 \\ 0 & 2 & 4 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \right\rangle,$$

and this is sufficient to confirm that A is diagonalizable. Further

$$\begin{aligned} & A^2 - (\lambda_1 + \lambda_2)A + \lambda_1\lambda_2 I \\ = \; & A^2 - 9A + 18I \\ = \; & \begin{pmatrix} 36 & 9 & 18 \\ 0 & 45 & 18 \\ 0 & -18 & 0 \end{pmatrix} - 9\begin{pmatrix} 6 & 1 & 2 \\ 0 & 7 & 2 \\ 0 & -2 & 2 \end{pmatrix} + \begin{pmatrix} 18 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 18 \end{pmatrix} = 0. \end{aligned}$$

So $A^2 = 9A - 18I$ can be written as a linear combination of A and I, and likewise

$$A^3 = 9A^2 - 18A = 81A - 180I$$

can also be written as such a linear combination. More generally (say using a proof by induction) we find that any polynomial in A can be written as a linear combination of A and I. However if $\operatorname{diag}(1,2,3) = \alpha A + \beta I$ for some α, β then, just looking at the diagonal entries, we'd have

$$6\alpha + \beta = 1, \qquad 7\alpha + \beta = 2, \qquad 2\alpha + \beta = 3,$$

and, with a quick check, we see this system is inconsistent. Hence $\operatorname{diag}(1,2,3)$ cannot be expressed as a polynomial in A and no such scalars $a_0, a_1, \ldots, a_n$ exist. ■

Example 57 *Determine x_n and y_n where $x_0 = 1$, $y_0 = 0$ and*

$$x_{n+1} = x_n - y_n \qquad \text{and} \qquad y_{n+1} = x_n + y_n \qquad \text{for } n \geqslant 0.$$

Solution. We can rewrite the two recurrence relations as a single recurrence relation involving a vector, namely

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

From Example 47 we have

$$P^{-1}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} P = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} \qquad \text{where} \quad P = \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}.$$

So

$$\begin{aligned} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^n &= P\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}^n P^{-1} \\ &= \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix} \begin{pmatrix} (1+i)^n & 0 \\ 0 & (1-i)^n \end{pmatrix} (-\tfrac{1}{2}) \begin{pmatrix} i & -1 \\ -1 & i \end{pmatrix} \\ &= \tfrac{1}{2}\begin{pmatrix} (1+i)^n + (1-i)^n & i(1-i)^n - i(1+i)^n \\ i(1-i)^n - i(1+i)^n & (1+i)^n + (1-i)^n \end{pmatrix} \\ &= \tfrac{1}{2}\begin{pmatrix} 2\operatorname{Re}(1+i)^n & 2\operatorname{Im}(1+i)^n \\ 2\operatorname{Im}(1+i)^n & 2\operatorname{Re}(1+i)^n \end{pmatrix}. \end{aligned}$$

By De Moivre's theorem, and noting $1+i=\sqrt{2}\operatorname{cis}(\pi/4)$, we have

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} \operatorname{Re}(1+i)^n & \operatorname{Im}(1+i)^n \\ \operatorname{Im}(1+i)^n & \operatorname{Re}(1+i)^n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \operatorname{Re}(1+i)^n \\ \operatorname{Im}(1+i)^n \end{pmatrix} = 2^{n/2} \begin{pmatrix} \cos(n\pi/4) \\ \sin(n\pi/4) \end{pmatrix}.$$

■

We briefly return to our first question from the start of the section: why might diagonalizability be a useful definition? We have seen that it can be computationally helpful, but representing a linear map by a diagonal matrix also helps us appreciate the effect of the linear map.

For each choice of basis of a finite dimensional vector space, a linear map is represented by a certain matrix. So a sensible question is: is there a preferential basis to best describe the linear map? Certainly if we can produce a diagonal matrix representative this will be a computational improvement but we will also better appreciate how the linear map transforms the vector space. To conclude, we recall that some matrices are not diagonalizable; this just invites the more refined question: into what preferred forms might we be able to change those matrices with a sensible choice of co-ordinates?