Further Mathematical Methods (Linear Algebra) 2002

Lecture 7: An Introduction To Spectral Theory

Here, we briefly revise some facts about eigenvalues and eigenvectors from MA100. As you know from this earlier course, these can be used to simplify certain geometric problems, but in this course, we shall use them in some other applications. The study of the eigenvalues and eigenvectors of a matrix is generally referred to as *spectral theory* since the set of eigenvalues of a matrix is sometimes called its *spectrum*.

7.1 Eigenvalues and Eigenvectors [in $\mathbb{R}^n$]

Consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$. Generally, there is no obvious geometrical relationship between the vectors $\mathbf{x}$ and $T(\mathbf{x})$ in $\mathbb{R}^n$, i.e. the vector $\mathbf{x}$ is transformed to a completely different vector $T(\mathbf{x})$ as both the magnitude and direction of $\mathbf{x}$ change under the action of T. However, *sometimes*, we find vectors $\mathbf{x}$ and $T(\mathbf{x})$ that have the same direction, i.e. the direction of $\mathbf{x}$ is *preserved* under the action of T if $\mathbf{x}$ is one of these 'special' vectors. So, such vectors, called the *eigenvectors* of the transformation T, are such that:

$$T(\mathbf{x}) = \lambda\mathbf{x},$$

for some $\lambda \in \mathbb{R}$ (since T is a transformation between two real spaces) and $\mathbf{x} \neq \mathbf{0}$ (since if $\mathbf{x} = \mathbf{0}$, then there is no direction to be preserved). Indeed, as we can represent the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ by an $n \times n$ matrix A where

$$T(\mathbf{x}) = \mathsf{A}\mathbf{x},$$

we say that

Definition 7.1 *If A is a [real] $n \times n$ matrix, then a non-zero vector $\mathbf{x} \in \mathbb{R}^n$ is called an eigenvector of A if*

$$\mathsf{A}\mathbf{x} = \lambda\mathbf{x},$$

for some $\lambda \in \mathbb{R}$. Here, λ is called an eigenvalue of A, and $\mathbf{x}$ is said to be an eigenvector corresponding to the eigenvalue λ.

To find the eigenvalues of an $n \times n$ matrix A, i.e. the scalars λ such that $\mathsf{A}\mathbf{x} = \lambda\mathbf{x}$ for some non-zero vector $\mathbf{x}$, we write this matrix equation as

$$\mathsf{A}\mathbf{x} = \lambda\mathsf{I}\mathbf{x},$$

or indeed,

$$(\mathsf{A} - \lambda\mathsf{I})\mathbf{x} = \mathbf{0}.$$

So, denoting the column vectors of the matrix $\mathsf{A} - \lambda\mathsf{I}$ by the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ we can write this matrix equation as a vector equation, i.e.

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbf{0},$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$. But, since we require that $\mathbf{x} \neq \mathbf{0}$, this means (by Definition 2.6) that the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$, and hence the column vectors of the matrix $\mathsf{A} - \lambda\mathsf{I}$, are linearly dependent. As such, the eigenvalues of the matrix are the scalars which make the column vectors of the matrix $\mathsf{A} - \lambda\mathsf{I}$ linearly dependent. Consequently, the eigenvalues of the matrix A are the scalars that make the matrix $\mathsf{A} - \lambda\mathsf{I}$ singular (i.e. not invertible) and as such they can be found by solving the determinant equation

$$\det(\mathsf{A} - \lambda\mathsf{I}) = 0.$$

Indeed, expanding this out, we have

Theorem 7.2 *If* A *is a [real]* $n \times n$ *matrix, then*

$$\det(\mathsf{A} - \lambda\mathsf{I}) = (-1)^n \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n,$$

and this is called the characteristic polynomial of A. *In particular, note that the coefficient of the* λ^n *term is* $(-1)^n$ *and* $c_n = \det(\mathsf{A})$.

Proof: See Problem Sheet 4, Question 11. ♠

and so, the eigenvalues of an $n \times n$ matrix A are just the roots of its characteristic polynomial. As such, A can have no more than n real eigenvalues.

Once we have found the eigenvalues of A, we can find the eigenvectors corresponding to each eigenvalue. Strictly speaking, this involves finding the set of all vectors which satisfy the matrix equation

$$(\mathsf{A} - \lambda\mathsf{I})\mathbf{x} = \mathbf{0},$$

for a given eigenvalue λ. Obviously, this set of vectors will be the null space of the matrix $\mathsf{A} - \lambda\mathsf{I}$, and we call this subspace of $\mathbb{R}^n$ the *eigenspace of* A corresponding to the eigenvalue λ. Indeed, any non-zero vector in this subspace will be an eigenvector of A corresponding to the eigenvalue λ. However, all we need to find to capture all of the information carried by these vectors is a basis for this subspace. So, when we talk of the eigenvectors of a matrix corresponding to some eigenvalue λ, we are just talking about a basis for the eigenspace.[1] Thus, to calculate the eigenvectors of the matrix A corresponding to the eigenvalue λ, we just need to find a basis for the solution space of the matrix equation above.

Indeed, we find that

Theorem 7.3 *Let* A *be a [real]* $n \times n$ *matrix. If* A *has* n *distinct [real] eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$, *then the eigenvectors corresponding to these eigenvalues form a linearly independent set.*

Proof: See Problem Sheet 4, Questions 2 and 8. ♠

Notice that in this case, where we have n distinct [real] eigenvalues, the eigenspace corresponding to each eigenvalue will be one-dimensional and together, the single basis vector which can be found for each eigenspace can be used to form a basis for $\mathbb{R}^n$. This prompts us to ask what happens when we do *not* have n distinct [real] eigenvalues.

For example: Consider the matrix given by

$$\mathsf{A} = \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix},$$

for some value of $a \in \mathbb{R}$. As discussed above, the eigenvalues of this matrix are the solutions to the determinant equation

$$\det(\mathsf{A} - \lambda\mathsf{I}) = 0,$$

that is, we have to solve

$$\begin{vmatrix} 1-\lambda & a \\ 1 & 1-\lambda \end{vmatrix} = 0 \implies (1-\lambda)^2 - a = 0.$$

But, clearly, the solutions of this equation are just $\lambda = 1 \pm \sqrt{a}$. From this, we can see that there are three possible cases:

- If $a > 0$, then A has two distinct real eigenvalues.
- If $a = 0$, then A has one [repeated] real eigenvalue, namely $\lambda = 1$.

[1] This is why the eigenvectors must be non-zero. (As any set that contains the null vector cannot be a basis since it will be linearly dependent. See the first part of Question 4 on Problem Sheet 1.)

- If $a < 0$, then A has two complex roots.[2]

(The second case will be discussed in Section 7.2 when we deal with repeated roots of the characteristic polynomial (i.e. 'multiplicitous' eigenvalues). Also, as we are only considering eigenvalues and eigenvectors in real space (i.e. $\mathbb{R}^2$) here we shall not discuss the third of these cases until Section 7.4.)

Now, to find the eigenvectors corresponding to the two distinct real eigenvalues (i.e. those where $a > 0$), we have to find a basis for the null space of the matrix $\mathsf{A} - \lambda\mathsf{I}$ for each value of λ. To do this, we just substitute the appropriate values of λ into the matrix equation

$$(\mathsf{A} - \lambda\mathsf{I})\mathbf{x} = \mathbf{0},$$

and solve for $\mathbf{x}$. Let us do this for the first case mentioned above:

When $a > 0$: In this case the eigenvalues are $\lambda = 1 \pm \sqrt{a}$ and we consider each of these in turn:

- If $\lambda = 1 + \sqrt{a}$, this substitution yields the matrix equation:

$$\begin{bmatrix} -\sqrt{a} & a \\ 1 & -\sqrt{a} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0},$$

and so, expanding out, we see that the components of these vectors must satisfy the simultaneous equations:

$$\begin{aligned} -\sqrt{a}x_1 + ax_2 &= 0 \\ x_1 - \sqrt{a}x_2 &= 0 \end{aligned}$$

But here, the first equation is just the second equation multiplied by the [non-zero] factor $-\sqrt{a}$, i.e. we effectively have just one equation (say, $x_1 - \sqrt{a}x_2 = 0$) relating two variables. Thus, one of these variables (say, x_2) must be free and so the null space of the matrix $\mathsf{A} - \lambda\mathsf{I}$ is given by all vectors of the form

$$\mathbf{x} = x_2 \begin{bmatrix} \sqrt{a} \\ 1 \end{bmatrix},$$

Consequently, we take $[\sqrt{a}, 1]^t$ to be the eigenvector corresponding to the eigenvalue $\lambda = 1 + \sqrt{a}$.[3]

- If $\lambda = 1 - \sqrt{a}$, this substitution yields the matrix equation:

$$\begin{bmatrix} \sqrt{a} & a \\ 1 & \sqrt{a} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0},$$

and so, expanding out, we see that the components of these vectors must satisfy the simultaneous equations:

$$\begin{aligned} \sqrt{a}x_1 + ax_2 &= 0 \\ x_1 + \sqrt{a}x_2 &= 0 \end{aligned}$$

But here, the first equation is just the second equation multiplied by the [non-zero] factor $\sqrt{a}$, i.e. we effectively have just one equation (say, $x_1 + \sqrt{a}x_2 = 0$) relating two variables. Thus, one of these variables (say, x_2) must be free and so the null space of the matrix $\mathsf{A} - \lambda\mathsf{I}$ is given by all vectors of the form

$$\mathbf{x} = x_2 \begin{bmatrix} -\sqrt{a} \\ 1 \end{bmatrix},$$

Consequently, we take $[-\sqrt{a}, 1]^t$ to be the eigenvector corresponding to the eigenvalue $\lambda = 1 - \sqrt{a}$.[4]

---

[2] Note that these form a complex conjugate pair. (This must be the case since the quadratic equation involved has real coefficients.)

[3] Since the set containing this vector, i.e. $\{[\sqrt{a}, 1]^t\}$, is a basis for the null space of $\mathsf{A} - \lambda\mathsf{I}$ when $\lambda = 1 + \sqrt{a}$.

[4] Since the set containing this vector, i.e. $\{[-\sqrt{a}, 1]^t\}$, is a basis for the null space of $\mathsf{A} - \lambda\mathsf{I}$ when $\lambda = 1 - \sqrt{a}$.

And so, we have found that $[\sqrt{a}, 1]^t$ and $[-\sqrt{a}, 1]^t$ are the eigenvectors corresponding to the eigenvalues $\lambda = 1 + \sqrt{a}$ and $\lambda = 1 - \sqrt{a}$ respectively in the case where $a > 0$. (Notice that we get two linearly independent eigenvectors since we have two distinct eigenvalues — see Theorem 7.3.) ♣

7.2 Multiplicity

If the characteristic polynomial of A does not have n distinct [real] roots, then we can also get repeated roots or complex roots. In this section, we consider what happens when [at least one] of the eigenvalues is a repeated root of the characteristic polynomial and this brings us onto the notion of *multiplicity*. Formally, we say:

Definition 7.4 *An eigenvalue λ' of a matrix* A *is said to be of multiplicity m if λ' is an m times repeated root of the characteristic polynomial of* A.[5]

Now, if we have a multiplicitous eigenvalue, the eigenspace of A corresponding to this eigenvalue can be more than one-dimensional. In particular, if λ is of multiplicity m, then the corresponding eigenspace can have a dimension of anything upto m depending on how many linearly independent eigenvectors we can find. Unfortunately, short of actually working out what the eigenvectors are, there is no easy way of determining how many of them there are! Thus, there is no guarantee that an eigenvalue of multiplicity m will have an m-dimensional eigenspace (i.e. m linearly independent eigenvectors) and this is why we can make no general claim about their linear independence.

For example: We now turn to the case where A has one [repeated] real eigenvalue (i.e. when $a = 0$) mentioned in the previous example. It is clear that in this case, the eigenvalue $\lambda = 1$ has multiplicity two[6] and we can now work out the eigenvectors corresponding to this eigenvalue using the method described above.

<u>When $a = 0$</u>: In this case the [repeated] eigenvalue is $\lambda = 1$ and considering this we have:

- For $\lambda = 1$, the substitution yields the matrix equation:

  $$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0},$$

  and so, expanding out, we see that the components of these vectors must satisfy the equation $x_1 = 0$. That is, we have just one equation relating two variables. Thus, one of these variables (namely, x_2) must be free and so the null space of the matrix $\mathsf{A} - \lambda\mathsf{I}$ is given by all vectors of the form

  $$\mathbf{x} = x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

  Consequently, we take $[0, 1]^t$ to be the eigenvector corresponding to the eigenvalue $\lambda = 1$.[7]

And so, we have found that $[0, 1]^t$ is the the eigenvector corresponding to the eigenvalue $\lambda = 1$ [in the case where $a = 0$].[8] (Notice that we have only one eigenvector here and so the dimension of the eigenspace corresponding to this [sole] eigenvalue is one.) ♣

[5]That is, the characteristic polynomial of A contains a factor of the form $(\lambda' - \lambda)^m$.

[6]Notice that when $a = 0$, the characteristic polynomial of the matrix A is given by

$$\det(\mathsf{A} - \lambda\mathsf{I}) = (1 - \lambda)^2,$$

as we should expect from the previous footnote.

[7]Since the set containing this vector, i.e. $\{[0, 1]^t\}$, is a basis for the null space of $\mathsf{A} - \lambda\mathsf{I}$ when $\lambda = 1$.

[8]This is what we should expect from setting $a = 0$ in the eigenvectors corresponding to the eigenvalues $\lambda = 1 \pm \sqrt{a}$ in the case where $a > 0$ discussed above.

7.3 Diagonalisation

We now briefly recall a useful technique which will allow us to simplify many matrix calculations in this course, namely how to *diagonalise* matrices. That is,

Definition 7.5 *A [real] $n \times n$ matrix A is called diagonalisable [over $\mathbb{R}$] iff there is a [real] invertible matrix P such that $\mathsf{P}^{-1}\mathsf{AP}$ is a [real] diagonal matrix.*

Indeed, it can be shown that

Theorem 7.6 *Let A be a [real] $n \times n$ matrix. A is diagonalisable [over $\mathbb{R}$] iff A has n [real] linearly independent eigenvectors.*

Proof: See Problem Sheet 4, Question 9. ♠

and so, if the matrix A has n distinct [real] eigenvalues then, by Theorem 7.3, it will be diagonalisable [over $\mathbb{R}$]. However, if there are multiplicitous eigenvalues, it may or may not be diagonalisable — it depends on whether we can find n linearly independent eigenvalues or not.[9]

It is actually straightforward to diagonalise a matrix A since the column vectors of the sought after invertible matrix P are just the eigenvectors of the matrix A.[10] Thus, if the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are the [linearly independent] eigenvectors corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ [including multiplicity] respectively, then

$$\mathsf{P} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix},$$

and this gives rise to the diagonal matrix D, i.e.

$$\mathsf{P}^{-1}\mathsf{AP} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \equiv \mathsf{D}.$$

Notice that the order in which the eigenvectors are put in to P determines the order in which the eigenvalues appear along the diagonal of D. Indeed, changing the order of the eigenvectors in P will change the order of the eigenvalues of D and so diagonalising a matrix does not lead to a unique diagonal matrix D.

For example: Looking at the matrix A in the earlier example, we can see that when $a > 0$ we have two linearly independent eigenvectors and so (by Theorem 7.6) it can be diagonalised. Following the prescription above, we let P be the matrix of eigenvectors, i.e.

$$\mathsf{P} = \begin{bmatrix} \sqrt{a} & -\sqrt{a} \\ 1 & 1 \end{bmatrix},$$

and as these vectors are linearly independent, $\det(\mathsf{P}) \neq 0$ and so this matrix is invertible. Further, if we calculate the matrix product $\mathsf{P}^{-1}\mathsf{AP}$, we find that

$$\begin{aligned} \mathsf{P}^{-1}\mathsf{AP} &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 1 & \sqrt{a} \\ -1 & \sqrt{a} \end{bmatrix} \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{a} & -\sqrt{a} \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 1 & \sqrt{a} \\ -1 & \sqrt{a} \end{bmatrix} \begin{bmatrix} \sqrt{a}+a & -\sqrt{a}+a \\ \sqrt{a}+1 & -\sqrt{a}+1 \end{bmatrix} \\ &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 2\sqrt{a}(1+\sqrt{a}) & 0 \\ 0 & 2\sqrt{a}(1-\sqrt{a}) \end{bmatrix} \\ \therefore \quad \mathsf{P}^{-1}\mathsf{AP} &= \begin{bmatrix} 1+\sqrt{a} & 0 \\ 0 & 1-\sqrt{a} \end{bmatrix} \end{aligned}$$

[9] Note: If some of the eigenvalues are complex, then (as you saw in MA100) the matrix is not disagonalisable over $\mathbb{R}$ since this requires that the eigenvalues (i.e. the entries in D — see below) are real numbers. (But, it may be diagonalisable over $\mathbb{C}$ — see the next section – if we can find n linearly independent eigenvectors corresponding to the n eigenvalues.)

[10] This is why we require that A has n linearly independent eigenvectors, as this guarantees that P is invertible.

and this is the required diagonal matrix D. (Notice that the entries of this matrix are the eigenvalues of A and that these appear in the same order as the eigenvectors corresponding to them appear in P.)

However, when $a = 0$ we only get one [linearly independent] eigenvector and so (by Theorem 7.6) the matrix A is *not* diagonalisable in this case.[11] ♣

7.4 Eigenvalues and Eigenvectors [in $\mathbb{C}^n$]

Presently, we will want to consider eigenvalues and eigenvectors of linear transformations which are defined over complex spaces such as $T : \mathbb{C}^n \to \mathbb{C}^n$. The need for such transformations can arise when we consider certain $n \times n$ matrices A representing a linear transformation T, for example:

- The matrix A could have complex entries and so it will map [possibly real vectors] to complex vectors, i.e. we could have $T : \mathbb{R}^n \to \mathbb{C}^n$.
- The matrix A may have real entries but these give rise to complex eigenvalues, i.e. the eigenvectors corresponding to the complex eigenvalues will be mapped to a complex space by T.

In these circumstances the results above need to be amended to take into account that the matrices, eigenvalues and eigenvectors can be complex. But, as we shall see, these modifications are straightforward as can be seen in Problem Sheet 4 where we make no assumptions about the set of scalars (i.e. $\mathbb{R}$ or $\mathbb{C}$) being used in the results of Questions 2, 8, 9, 10 and 11.

For example: We now turn to the case where A has two complex eigenvalues (i.e. when $a < 0$) mentioned in the first example. Now, to find the eigenvectors corresponding to these eigenvalues we use the method described in Section 7.1 (despite the fact that they are complex numbers) and so we just have to find a basis for the null space of the matrix $\mathsf{A} - \lambda \mathsf{I}$ for each value of λ. As before, we do this by substituting the appropriate values of λ into the matrix equation

$$(\mathsf{A} - \lambda \mathsf{I})\mathbf{x} = \mathbf{0},$$

and solve for $\mathbf{x}$. So, doing this, we find that:

When $a < 0$: In this case the eigenvalues are $\lambda = 1 \pm \sqrt{a}$ and we consider each of these in turn:

- If $\lambda = 1 + \sqrt{a}$, this substitution yields the matrix equation:

$$\begin{bmatrix} -\sqrt{a} & a \\ 1 & -\sqrt{a} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0},$$

  and so, expanding out, we see that the components of these vectors must satisfy the simultaneous equations:

$$\begin{aligned} -\sqrt{a}x_1 + ax_2 &= 0 \\ x_1 - \sqrt{a}x_2 &= 0 \end{aligned}$$

  But here, the first equation is just the second equation multiplied by the [non-zero] factor $-\sqrt{a}$, i.e. we effectively have just one equation (say, $x_1 - \sqrt{a}x_2 = 0$) relating two variables. Thus, one of these variables (say, x_2) must be free and so the null space of the matrix $\mathsf{A} - \lambda \mathsf{I}$ is given by all vectors of the form

$$\mathbf{x} = x_2 \begin{bmatrix} \sqrt{a} \\ 1 \end{bmatrix},$$

  Consequently, we take $[\sqrt{a}, 1]^t$ to be the eigenvector corresponding to the eigenvalue $\lambda = 1 + \sqrt{a}$.[12]

[11] Notice that A fails to be diagonalisable in this case because we can't find two linearly independent eigenvectors to construct an *invertible* matrix P.

[12] Since the set containing this vector, i.e. $\{[\sqrt{a}, 1]^t\}$, is a basis for the null space of $\mathsf{A} - \lambda \mathsf{I}$ when $\lambda = 1 + \sqrt{a}$.

- If $\lambda = 1 - \sqrt{a}$, this substitution yields the matrix equation:

$$\begin{bmatrix} \sqrt{a} & a \\ 1 & \sqrt{a} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0},$$

and so, expanding out, we see that the components of these vectors must satisfy the simultaneous equations:

$$\begin{aligned} \sqrt{a}x_1 + ax_2 &= 0 \\ x_1 + \sqrt{a}x_2 &= 0 \end{aligned}$$

But here, the first equation is just the second equation multiplied by the [non-zero] factor $\sqrt{a}$, i.e. we effectively have just one equation (say, $x_1 + \sqrt{a}x_2 = 0$) relating two variables. Thus, one of these variables (say, x_2) must be free and so the null space of the matrix $\mathsf{A} - \lambda\mathsf{I}$ is given by all vectors of the form

$$\mathbf{x} = x_2 \begin{bmatrix} -\sqrt{a} \\ 1 \end{bmatrix},$$

Consequently, we take $[-\sqrt{a}, 1]^t$ to be the eigenvector corresponding to the eigenvalue $\lambda = 1 - \sqrt{a}$.[13]

And so, unsurprisingly, we have found that $[\sqrt{a}, 1]^t$ and $[-\sqrt{a}, 1]^t$ are the [now complex] eigenvectors corresponding to the [complex] eigenvalues $\lambda = 1 + \sqrt{a}$ and $\lambda = 1 - \sqrt{a}$ respectively in the case where $a < 0$. (Notice that we get two linearly independent eigenvectors since we have two distinct eigenvalues — see Theorem 7.3.)

Further, since we have two linearly independent eigenvectors (by Theorem 7.6) we can diagonalise A. So, following the prescription in Section 7.3, we let P be the [complex] matrix of eigenvectors, i.e.

$$\mathsf{P} = \begin{bmatrix} \sqrt{a} & -\sqrt{a} \\ 1 & 1 \end{bmatrix},$$

and as these vectors are linearly independent, $\det(\mathsf{P}) \neq 0$ and so this matrix is invertible. Further, if we calculate the matrix product $\mathsf{P}^{-1}\mathsf{AP}$, we find that

$$\begin{aligned} \mathsf{P}^{-1}\mathsf{AP} &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 1 & \sqrt{a} \\ -1 & \sqrt{a} \end{bmatrix} \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{a} & -\sqrt{a} \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 1 & \sqrt{a} \\ -1 & \sqrt{a} \end{bmatrix} \begin{bmatrix} \sqrt{a} + a & -\sqrt{a} + a \\ \sqrt{a} + 1 & -\sqrt{a} + 1 \end{bmatrix} \\ &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 2\sqrt{a}(1 + \sqrt{a}) & 0 \\ 0 & 2\sqrt{a}(1 - \sqrt{a}) \end{bmatrix} \\ \therefore \quad \mathsf{P}^{-1}\mathsf{AP} &= \begin{bmatrix} 1 + \sqrt{a} & 0 \\ 0 & 1 - \sqrt{a} \end{bmatrix}, \end{aligned}$$

and this is the required diagonal matrix D. (Notice that the entries of this [complex] matrix are the [complex] eigenvalues of A and that these appear in the same order as the [complex] eigenvectors corresponding to them appear in P.)

So, we can see that if the eigenvalues of a matrix are complex, the same methods for finding the eigenvalues and eigenvectors apply, as does the method for diagonalising the matrix. However, we are now working in a complex space (in this example, $\mathbb{C}^2$) when we use them. ♣

[13] Since the set containing this vector, i.e. $\{[-\sqrt{a}, 1]^t\}$, is a basis for the null space of $\mathsf{A} - \lambda\mathsf{I}$ when $\lambda = 1 - \sqrt{a}$.