Lecture 2

2.1 Diagonal Matrices

In the first lecture, the focus was upon general square matrices. This time we specialise by looking at a category of square matrix that has interesting properties; the diagonal matrix.

Definition : A Diagonal Matrix

A diagonal matrix is a square matrix in which all elements not on the leading diagonal are zero.

It is of the form $\begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & a_{nn} \end{pmatrix}$ where $a_{ij} = 0$ if $i \neq j$

The initial interest in diagonal matrices arises from the following simple observation;

For a diagonal matrix $\mathbf{D} = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & a_{nn} \end{pmatrix}$, $\mathbf{D}^k = \begin{pmatrix} a_{11}^k & 0 & 0 & \dots & 0 \\ 0 & a_{22}^k & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & a_{nn}^k \end{pmatrix}$

So, for example, $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^4 = \begin{pmatrix} 2^4 & 0 \\ 0 & 3^4 \end{pmatrix} = \begin{pmatrix} 16 & 0 \\ 0 & 81 \end{pmatrix}$

By itself, this is unremarkable. However, many other matrices, **A**, can be written in the form $\mathbf{A} = \mathbf{P}\ \mathbf{D}\ \mathbf{P}^{-1}$ for some matrix **P** and a diagonal matrix **D**. This suggests a possible shortcut to working out $\mathbf{A}^k$ because,

$$\begin{aligned} \mathbf{A}^k &= \left(\mathbf{PD}\,\mathbf{P}^{-1}\right)^k \\ &= \left(\mathbf{PD}\,\mathbf{P}^{-1}\right)\left(\mathbf{PD}\,\mathbf{P}^{-1}\right)\dots\left(\mathbf{PD}\,\mathbf{P}^{-1}\right) \\ &= \mathbf{PD}\left(\mathbf{P}^{-1}\,\mathbf{P}\right)\mathbf{D}\left(\mathbf{P}^{-1}\,\mathbf{P}\right)\dots\left(\mathbf{P}^{-1}\,\mathbf{P}\right)\mathbf{D}\,\mathbf{P}^{-1} \\ &= \mathbf{P}\,\mathbf{D}^k\,\mathbf{P}^{-1} \end{aligned}$$

For this to be worthwhile, given a matrix **A**, it would need to be relatively straight forward to find the corresponding matrices **P**, **D** and $\mathbf{P}^{-1}$ ahead of using the fact that $\mathbf{A}^k = \mathbf{P}\,\mathbf{D}^k\,\mathbf{P}^{-1}$

2.2 How to Find P and D

Given a $n \times n$ square matrix **A** that is diagonalizable, **D** and **P** are found by;

Step 1: Find the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ of the matrix **A**
(If there are n distinct eigenvalues, **A** is diagonalizable)

Step 2: $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & 0 & ... & 0 \\ 0 & \lambda_2 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & 0 & 0 & \lambda_n \end{pmatrix}$ where $a_{ij} = 0$ if $i \neq j$, $a_{ij} = \lambda_i$ if $i = j$

Step 3: Find the corresponding eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n$ of the matrix A

Step 4: $\mathbf{P} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & ... & \boldsymbol{v}_n \end{pmatrix}$

2.2.1 Example

Question: Find the matrices **P** and **D** given that $\mathbf{A} = \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}$ and hence $\mathbf{A}^{100}$

Answer:

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

$$det\begin{pmatrix} 3 - \lambda & -1 \\ 2 & -\lambda \end{pmatrix} = 0$$

$$(3 - \lambda)(-\lambda) + 2 = 0$$

$$\lambda^2 - 3\lambda + 2 = 0$$

$$(\lambda - 1)(\lambda - 2) = 0$$

$$\lambda_1 = 1 \text{ or } \lambda_2 = 2$$

$$\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

For $\lambda_1 = 1 : \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 1\begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow y = 2x \Rightarrow \boldsymbol{v}_1 = k_1\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \ k_1 \in \mathbb{R}$

For $\lambda_2 = 2 : \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow y = x \Rightarrow \boldsymbol{v}_2 = k_2\begin{pmatrix} 1 \\ 1 \end{pmatrix}, k_2 \in \mathbb{R}$

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \text{ from which } \mathbf{P}^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

$$\mathbf{A}^k = \mathbf{P}\,\mathbf{D}^k\,\mathbf{P}^{-1}$$

$$\mathbf{A}^{100} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^{100}\begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2^{100} \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 2^{101} - 1 & 1 - 2^{100} \\ 2^{101} - 2 & 2 - 2^{100} \end{pmatrix}$$

2.3 Understanding Diagonalization

To recap, so far it has been observed that $n \times n$ matrices with n distinct eigenvalues, **A**, can be written in the form $\mathbf{A} = \mathbf{P\,D\,P}^{-1}$ where **D** is a diagonal matrix. The motivation to do this was to find a shortcut method to calculate powers of **A** by using the deduced result that,

$$\mathbf{A}^k = \mathbf{P\,D}^k\,\mathbf{P}^{-1}$$

However, the fact that **A** has this matrix **D** associated with it is interesting in its own right. To better understand this association a specific example will be considered, this time with an eye upon the geometry it represents.

Our fresh example will feature the matrix $\mathbf{A} = \begin{pmatrix} 11 & -2 \\ -2 & 14 \end{pmatrix}$

Its eigenvalues are found in the established fashion as follows;

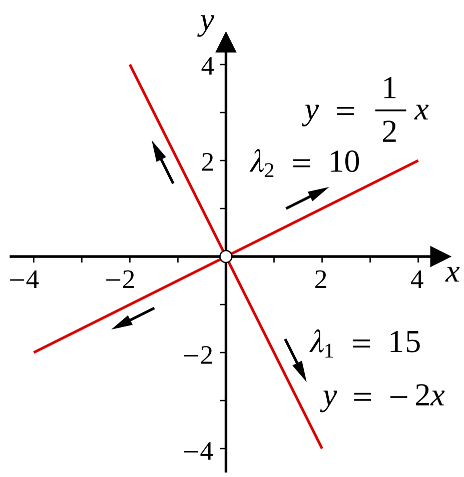
$$\begin{aligned} det(\mathbf{A} - \lambda\mathbf{I}) &= 0 \\ det\begin{pmatrix} 11-\lambda & -2 \\ -2 & 14-\lambda \end{pmatrix} &= 0 \\ (11-\lambda)(14-\lambda) - 4 &= 0 \\ \lambda^2 - 25\lambda + 150 &= 0 \\ (\lambda - 15)(\lambda - 10) &= 0 \\ \lambda_1 = 15 \text{ or } \lambda_2 &= 10 \end{aligned}$$

In deducing the eigenvectors, this time they will be normalized. By making them have a length of unity, the matrix **P** and its inverse will be easier to interpret geometrically.

For $\lambda_1 = 15$: $\begin{pmatrix} 11 & -2 \\ -2 & 14 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 15\begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow y = -2x \Rightarrow \|\mathbf{v}_1\| = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -2 \end{pmatrix}$

For $\lambda_2 = 10$: $\begin{pmatrix} 11 & -2 \\ -2 & 14 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 10\begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow y = \frac{1}{2}x \Rightarrow \|\mathbf{v}_2\| = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Geometrically, the following diagram shows the two invariant lines along which the eigenvalues give the length scaling factor.



The matrices $\mathbf{P}$, $\mathbf{D}$ and $\mathbf{P}^{-1}$ can now be written down following the steps described previously,

$$\mathbf{P} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}, \ \mathbf{D} = \begin{pmatrix} 15 & 0 \\ 0 & 10 \end{pmatrix} \text{ and } \mathbf{P}^{-1} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

Geometrically, D is a scaling matrix that scales with a length scale factor of 15 along the x-axis and 10 along the y, easily seen from the following calculation,

$$\begin{pmatrix} 15 & 0 \\ 0 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 15x \\ 10y \end{pmatrix}$$

A rotation matrix has the form $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ and, amazingly, both $\mathbf{P}$ and $\mathbf{P}^{-1}$ are such matrices. For $\mathbf{P}^{-1}$, $\cos\theta = \frac{1}{\sqrt{5}}$ and $\sin\theta = \frac{2}{\sqrt{5}}$ and the angle of (anticlockwise) rotation is given by

$$\begin{aligned} \tan\theta &= \frac{\sin\theta}{\cos\theta} \\ &= \frac{2}{1} \\ \theta &= 63.4° \end{aligned}$$

A similar calculation shows that $\mathbf{P}$ is a same angle rotation but in the opposite direction, clockwise.

The matrix decomposition $\mathbf{A} = \mathbf{P}\,\mathbf{D}\,\mathbf{P}^{-1}$ in this example is thus saying that,

- A scaling of 15 along $y = -2x$ and 10 along $y = \frac{1}{2}x$

can also be achieved by,

- Rotating the lines $y = -2x$ and $y = \frac{1}{2}x$ by $63.4°$ $\quad \mathbf{P}^{-1}$

  (such that they lie along the x-axis and the y-axis respectively)
- Scaling by 15 along the x-axis and by 10 along the y-axis $\quad \mathbf{D}$
- Rotating the x and y-axis back onto $y = -2x$ and $y = \frac{1}{2}x$ $\quad \mathbf{P}$

Notice that the matrices $\mathbf{A}$ and $\mathbf{D}$ have the same characteristic equation.

Our example covered a most useful special case (that will be investigated further next) in which the invariant lines along which the eigenvectors and eigenvalues acted were mutually perpendicular. In the more general case, where this is not so, the matrix $\mathbf{P}$ has to do more than just rotate; it has to adjust the angle between the invariant lines to map them onto the x and y axes. This makes the geometric interpretation more involved. Exploring such is left as an exercise for the interested reader !

2.4 The Orthonormal Matrix

The section 2.3 example involved a matrix **A** that was orthogonal. To explain what that is, a detour is required involving the transpose matrix, the symmetric matrix, orthogonal vectors and orthonormal vectors. However, orthogonal matrices have a marvellous property that will make the journey worthwhile.

Transpose of a Matrix → Symmetric Matrix → Orthonormal Vectors → Orthogonal Matrix

---

Definition : The Transpose of an $n \times n$ Matrix

Given the matrix **M** $= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$ the transpose of matrix **M**

is denoted $\mathbf{M}^{\mathrm{T}}$ and is formed by an interchange of rows and columns.

Thus,
$$\mathbf{M}^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

For example, $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}$

Definition : A Symmetric Matrix

A matrix, **M**, is symmetric if $\mathbf{M} = \mathbf{M}^{\mathrm{T}}$. Such matrices are readily recognised for their elements are symmetric with respect to the leading diagonal.

For example, $\begin{pmatrix} a & b & c \\ b & e & f \\ c & f & g \end{pmatrix}$ is symmetric (about its leading diagonal shown red)

Two vectors are described as being orthogonal if they are perpendicular to each other. One way to test two (non-zero) vectors to determine if they are orthogonal is to calculate their scalar (dot) product. If that is zero then the the vectors are orthogonal, otherwise not. A set of vectors $\{ \boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n \}$ are mutually orthogonal if every pair of vectors is orthogonal. In other words if,

$$\boldsymbol{v}_i \bullet \boldsymbol{v}_j = 0 \text{ for all } i \neq j$$

Example: To show the vectors $\begin{pmatrix} \sqrt{2} \\ 0 \\ -\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$ are mutually orthogonal work through the three possible scalar products showing they are all zero.

$$\begin{pmatrix} \sqrt{2} \\ 0 \\ -\sqrt{2} \end{pmatrix} \bullet \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = \sqrt{2} \times 1 + 0 \times \sqrt{2} + \left(-\sqrt{2}\right) \times 1 = 0$$

$$\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = 1 \times 1 + \sqrt{2} \times \left(-\sqrt{2}\right) + 1 \times 1 = 0$$

$$\begin{pmatrix} \sqrt{2} \\ 0 \\ -\sqrt{2} \end{pmatrix} \bullet \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = \sqrt{2} \times 1 + 0 \times \left(-\sqrt{2}\right) + \left(-\sqrt{2}\right) \times 1 = 0$$

As all scalar products are zero the three vectors are mutually orthogonal. ☐

If a set of mutually orthogonal vectors are normalized they are termed a set of orthonormal vectors. Normalizing a vector means keeping its direction but scaling it so that it has a length of 1. This is done by dividing the vector by its magnitude,

$$\|\boldsymbol{v}\| = \frac{\boldsymbol{v}}{|\boldsymbol{v}|}$$

where $\|\boldsymbol{v}\|$ denotes a vector $\boldsymbol{v}$, normalized, and $|\boldsymbol{v}|$ is the vector's magnitude.

The vectors of the previous example, normalized, become,

$$\left\|\begin{pmatrix} \sqrt{2} \\ 0 \\ -\sqrt{2} \end{pmatrix}\right\| = \frac{1}{2}\begin{pmatrix} \sqrt{2} \\ 0 \\ -\sqrt{2} \end{pmatrix}, \quad \left\|\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}\right\| = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \quad \left\|\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}\right\| = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$$

Definition : An Orthogonal Matrix

Given a set of n orthonormal n-dimensional vectors, $\{v_1, v_2, ..., v_n\}$, the $n \times n$ square matrix $\mathbf{M} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \dots & \boldsymbol{v}_n \end{pmatrix}$ with real entries is termed an orthogonal matrix. An orthogonal matrix, M, has the marvellous property that,

$$\mathbf{M}^{\mathrm{T}}\mathbf{M} = \mathbf{I} = \mathbf{M}\,\mathbf{M}^{\mathrm{T}} \Leftrightarrow \mathbf{M}^{-1} = \mathbf{M}^{\mathrm{T}}$$

Working out the inverse of an $n \times n$ matrix for $n > 2$ can be hard work but for an orthogonal matrix is easy as it is simply the transpose.

Thus, building on the previous example, the following matrix, **M**, is orthogonal and has an inverse that is simply its transpose.

$$\mathbf{M} = \frac{1}{2}\begin{pmatrix} \sqrt{2} & 1 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 1 & 1 \end{pmatrix}, \quad \mathbf{M}^{-1} = \mathbf{M}^{\mathrm{T}} = \frac{1}{2}\begin{pmatrix} \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \\ 1 & -\sqrt{2} & 1 \end{pmatrix}$$

The reader may like to check that, for this example, $\mathbf{M}\,\mathbf{M}^{\mathrm{T}} = \mathbf{I}$

2.5 Exercise

Any solution based entirely on graphical
or numerical methods is not acceptable
Marks Available : 50

Question 1

Decompose the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ into the form $\mathbf{P}\,\mathbf{D}\,\mathbf{P}^{-1}$ where $\mathbf{D}$ is a diagonal matrix and hence determine $\mathbf{A}^{100}$ leaving the entries in index form.

[6 marks]

Question 2

A-Level Examination Question from May 2018, Paper FPM2, Q4 (Edexcel)

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$$

Find a matrix **P** and a diagonal matrix **D** such that $\mathbf{D} = \mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}$

[7 marks]

Question 3

A-Level Sample Assessment Materials from 2018, Paper FPM2, Q2 (Edexcel)

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 5 & 4 \\ 4 & 4 & 3 \end{pmatrix}$$

(a) Verify $\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$ is an eigenvector of **A** and find the corresponding eigenvalue.

[3 marks]

(b) Show 9 is an eigenvalue of **A** and find the corresponding eigenvector.

[5 marks]

Given that a third eigenvector of **A** is $\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$

(c) write down a matrix **P** and a diagonal matrix **D** such that $\mathbf{P}^{-1}\mathbf{A}\,\mathbf{P} = \mathbf{D}$

[4 marks]

Question 4

The matrix $\mathbf{A} = \begin{pmatrix} 7 & 0 & -2 \\ 0 & 5 & -2 \\ -2 & -2 & 6 \end{pmatrix}$

(a) Given 9 is an eigenvalue, find the other two eigenvalues of **A**.

[5 marks]

(b) Find the eigenvectors corresponding to the eigenvalues of **A**.

[5 marks]

(c) Show the eigenvectors found in part (b) are mutually perpendicular.

[2 marks]

(d) Find a matrix **P** and a diagonal matrix **D** such that $\mathbf{P}^{T}\ \mathbf{A}\ \mathbf{P} = \mathbf{D}$

[2 marks]

Question 5

Open University (UK) Specimen Examination Question from 1994, M203, Q4

The linear transformation $f : \mathbb{R}^3 \to \mathbb{R}^3$ is represented by the matrix,

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -3 & 4 \\ 0 & -2 & 3 \end{pmatrix}$$

(a) Find the eigenvalues of **A** with respect to the standard basis.

[2 marks]

(b) Find a basis for $\mathbb{R}^3$ consisting of eigenvectors of **A**.
(That is, find corresponding eigenvectors)

[4 marks]

Question 6

Princeton University Examination Question, Linear Algebra Paper, 2009

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

(a) Find the characteristic polynomial and all the eigenvalues (real and complex) of **A**.

[4 marks]

(b) Giving a reason, state if **A** is diagonizable over the complex numbers.

[2 mark]

(c) Calculate $\mathbf{A}^{2009}$

[4 marks]