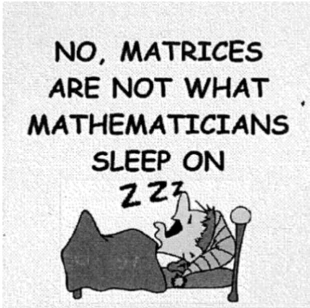
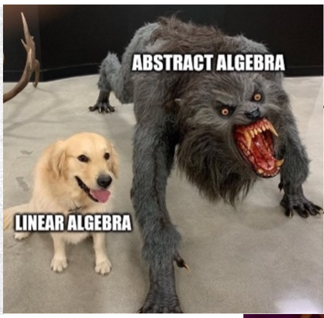
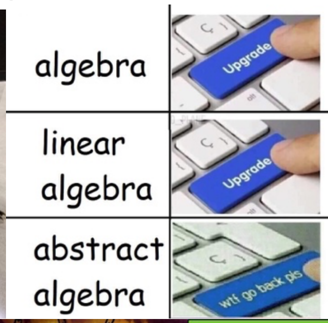
Linear Algebra: Eigenvectors, Eigenvalues and Diagonalisation







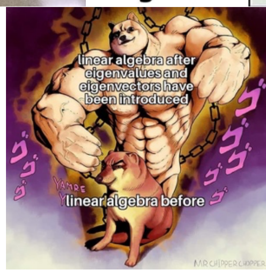


Table of Contents

1 Intuition

Personally, I feel that intuition isn't something which is easily explained. Intuition in mathematics is synonymous with experience and you gain intuition by working numerous examples, but let's try and get some intuition first on eigenvalue and eigenvectors.

In Life:

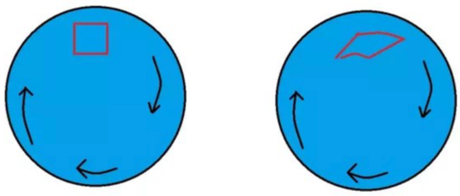
You see a beautiful smart girl/boy and your heart starts to beat a bit faster. However, there is that one person who just makes your heart beat like crazy. You are extremely happy to see that person and be with them. That person is your eigenvector and the amount of increase in happiness is the eigenvalue.

Objects:

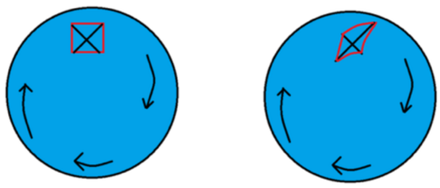
Consider a glass of water.

Now look at it from the top view.

Take a spoon and start mixing the water in a rotating clockwise fashion. After that, take out the spoon whilst the water is still rotating in the clockwise direction. Now, let's put a square made out of perfectly elastic rubber as shown below.



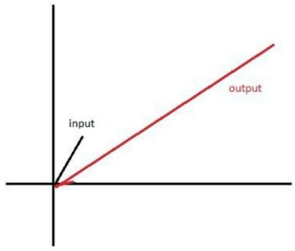
As the water on the edge is rotating faster than the water at the centre, the elastic will come in position 2 as shown. What happened?
The rubber square sides got stretched. But, if you look at the diagonals of the square, as shown below, one of the diagonals is stretching. However, both the diagonals have the same direction, even after getting stretched.



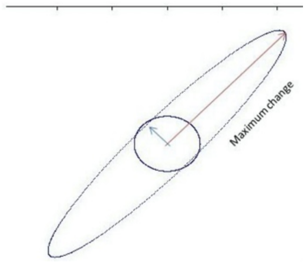
In linear algebra, an eigenvector or characteristic vector of a linear transformation is a non-zero vector whose direction does not change when that linear transformation is applied to it. So, for our case, the diagonals were applied a linear stretching/contracting transformation by the rotating water, the direction remained constant. So, the diagonals are out eigenvectors. And how much did they expand/contract? The value by they are doing so is called an eigenvalue.

Algebraically and Graphically:

Consider a matrix which is your system. $s = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$. The input can be any vector for example $X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. When operated on X, your system will give an output Y as Y= $\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix}=\begin{pmatrix} 8 \\ 9 \end{pmatrix}$. We can see in a picture that the initial vector (black) is transformed to another vector (red).



Now lets do it for all the points on the circle which are at equal distance from the centre so we can see the effect.



As seen, there is one particular vector, whose change in length is maximum. That change in length is the eigenvalue and vector is the eigenvector.

Conically:
Another way to explain this would be to go back to the old ellipsoid. The problem set being take a bunch of data points and find the ellipsoid that best fits the data. The 1st eigenvector will be the semi-major axis and the value is the magnitude and the 2nd one is the semi major axis. From there it is easy to explain how to carry it forward for n dimensions.

Statistically:
Another is to use the notion of "explained variance". Suppose you have points in n-dimensions. What would be the 1st vector that one should use to describe the points where minimizing the error (In an L2 sense)? What would be the second one, etc? These are the corresponding eigenvectors/values.

More abstractly:
First let us think what a square matrix does to a vector. Consider a matrix $A \in \mathbb{R}^{n \times n}$ (sometimes written $M_{n \times n}(\mathbb{R})$). Let us see what the matrix A acting on a vector x does to this vector. By action, we mean multiplication i.e. we get a new vector $y = A\boldsymbol{x}$. The matrix acting on a vector x does two things to the vector x.

1. It scales the vector

   $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ stretch by scale factor a parallel to x axis

   $\begin{pmatrix} 0 & 0 \\ 0 & a \end{pmatrix}$ stretch by scale factor a parallel to y axis

   $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ Enlargements by scale factor a

2. It rotates the vector

   $\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$ by an angle θ anticlockwise about origin

However, for any matrix A, there are some favoured vectors/directions. When the matrix acts on these favoured vectors, the action essentially results in just scaling the vector. There is no rotation. These favoured vectors are precisely the eigenvectors and the amount by which each of these favoured vectors stretches or compresses is the eigenvalue.

So why are these eigenvectors and eigenvalues important? Consider the eigenvector corresponding to the maximum (absolute) eigenvalue. If we take a vector along this eigenvector, then the action of the matrix is maximum. **No other vector when acted by this matrix will get stretched as much as this eigenvector**. Hence, if a vector were to lie "close" to this eigen direction, then the "effect" of action by this matrix will be "large" i.e. the action by this matrix results in "large" response for this vector. The effect of the action by this matrix is high for large (absolute) eigenvalues and less for small (absolute) eigenvalues. Hence, the directions/vectors along which this action is high are called the principal directions or principal eigenvectors. The corresponding eigenvalues are called the principal values.

The eigenvectors are the "axes" of the transformation represented by the matrix. Consider spinning a globe (the universe of vectors): every location faces a new direction, except the poles. The eigenvalue is the amount the eigenvector is scaled up or down when going through the matrix. Eigenvalues are special numbers associated with a matrix and eigenvectors are special vectors.
A matrix 'A' acts on vectors v like a function does, with input v and output Av. Eigenvectors are vectors for which Av is parallel to v. In other words: $av = \lambda v$ where v is an eigenvector and λ is an eigenvalue.

So, eigenvectors and eigenvalues are not something you use in a matrix. They are something a matrix has. An eigenvector is a vector which, multiplied by the matrix does not change direction. It may change length. The eigenvalue of an eigenvector is the magnitude by which the vector length changed. The rotation matrix rotates all vectors except those that match its rotation axis. These vectors are its eigenvectors and their eigenvalues are 1. Another example could be a scaling matrix which doubles the length of any vector. All vectors are its eigenvectors and their eigenvalues are equal to 2.

Eigenvalues play a key role in diagonalising matrices. The eigenvalues are the values that occur on the main diagonal in such an approximation. Not every matrix can be diagonalised i.e. that is to say no equal to a diagonal matrix after a basis change.

2 Definitions

Matrices have an algebra associated with them. You can multiply them together, add them, invert them and transpose them. The elements can be complex values, not just reals.

One special property are the "eigenvalues" and "eigenvectors" of a matrix. Eigenvectors and eigenvalues are linked and dependent on the matrix. So, by definition, the eigenvectors and their corresponding eigenvalues for a given matrix A are the solutions to:

$$A\boldsymbol{v} = \lambda\boldsymbol{v}$$

Where λ , which is a scalar, is an eigenvalue and $\boldsymbol{v}$ is an eigenvector. Again, the eigenvectors & eigenvalues of a system are of significance.

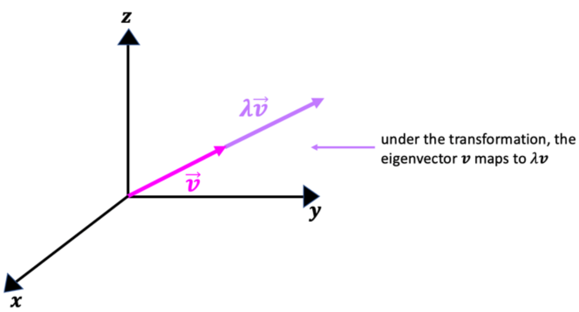
First let us think what a square matrix does to a vector. Consider a matrix $A \in \mathbb{R}^{n\times n}$ (you might also see this notation as $M_{n\times n}(\mathbb{R})$). Let us see what the matrix A acting on a vector $\boldsymbol{v}$ does to this vector. By action, we mean multiplication i.e. we get a new vector $y = A\boldsymbol{v}$.
The matrix A acting on a vector $\boldsymbol{v}$ does two things to the vector $\boldsymbol{v}$.

1. It scales the vector
2. It rotates the vector

However, for any matrix A, there are some favoured vectors/directions. When the matrix acts on these favoured vectors, the action essentially results in just scaling the vector. There is no rotation. These favoured vectors are precisely the eigenvectors and the amount by which each of these favoured vectors stretches or compresses is the eigenvalue.

So why are these eigenvectors and eigenvalues important? Consider the eigenvector corresponding to the maximum (absolute) eigenvalue. If we take a vector along this eigenvector, then the action of the matrix is maximum. **No other vector when acted by this matrix will get stretched as much as this eigenvector**. Hence, if a vector were to lie "close" to this eigen direction, then the "effect" of action by this matrix will be "large" i.e. the action by this matrix results in "large" response for this vector. The effect of the action by this matrix is high for large (absolute) eigenvalues and less for small (absolute) eigenvalues. Hence, the directions/vectors along which this action is high are called the principal directions or principal eigenvectors. The corresponding eigenvalues are called the principal values.

The image of an eigenvector $\boldsymbol{v}$ under a linear transformation A has the same direction as v, but may have a different magnitude. The eigenvalue can be interpreted as the magnification factor of the eigenvector under the transformation.



If v is an eigenvector of a matrix A representing a linear transformation, then the straight line that passes through the origin in the direction of v is an invariant line under than transformation. If the corresponding eigenvalue is 1, then every point on this line is an invariant point.

Linear Algebra can get very abstract in nature!

2.1.1 Characteristic Polynomial

If **v** is an eigenvector of the matrix A, then by definition

$$\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v}$$

This is the same as

$$\boldsymbol{A}\mathbf{v} = \lambda\boldsymbol{I}\mathbf{v} \text{ where } \boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ for a 2x2 matrix}$$

Re-arranging

$$\boldsymbol{A}\mathbf{v} - \lambda\boldsymbol{I}\mathbf{v} = \mathbf{0}$$

factoring out $\boldsymbol{v}$

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

By definition $\boldsymbol{v}$ is non-zero, the matrix $(\mathbf{A}-\lambda\boldsymbol{I})$ is singular and has determinant zero, so det $(\mathbf{A}-\lambda\boldsymbol{I})=0$
Note: $\det(\mathbf{A}-\lambda I)=\mathbf{0}$ is the same as $\det(\lambda-\boldsymbol{AI})=\mathbf{0}$ since we could have re-arranged as $\lambda\boldsymbol{I}\mathbf{v}-\boldsymbol{Av}=\mathbf{0}$ instead

$\det(\lambda-\boldsymbol{AI})$ is known as the characteristic polynomial/equation

Sometimes the characteristic equation has a single repeated or no real solutions. This leads to either repeated eigenvalues or complex eigenvalues. Note that:

- For a 2x2 matrix, the characteristic equation is a quadratic
- A 3x3 matrix will always have at least one real eigenvalue since a cubic equation always has at least one real solution
- The dimension of the matrix will be the power of characteristic polynomial.

Find the characteristic equation of the matrix $A=\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}$

$\mathrm{A}-\lambda I=\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}-\lambda\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}-\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}=\begin{pmatrix}2-\lambda & 5\\ -1 & -4-\lambda\end{pmatrix}$
$\det(\mathrm{A}-\lambda I)=\begin{vmatrix}2-\lambda & 5\\ -1 & -4-\lambda\end{vmatrix}=(2-\lambda)(-4-\lambda)-5(-1)=\lambda^2+2\lambda-3$

characteristic equation: $\lambda^2+2\lambda-3$

2.1.2 Eigenvectors and Eigenvalues

So far we have learnt that

$\mathbf{Av}=\lambda\mathbf{v}$ where λ =eigenvector and $\mathbf{v}$=eigenvector

$$\mathbf{Av}=\lambda\mathbf{v}\Leftrightarrow \boldsymbol{A}\mathbf{v}-\lambda\boldsymbol{I}\mathbf{v}=\mathbf{0}\Leftrightarrow(\mathbf{A}-\lambda\boldsymbol{I})\mathbf{v}=\mathbf{0}$$

Steps to find Eigen**values**:
Step 1: Find $\mathrm{A}-\lambda I$
Step 2: Set characteristic polynomial/equation which is $\det(\mathbf{A}-\lambda\boldsymbol{I})$ or $\det(\lambda-\boldsymbol{AI})$ equal to zero 0, where **A** is the matrix and $\boldsymbol{I}$ is the identity matrix.
Step 3: Solve this and you get the eigenvalue(s) λ.

Steps to find Eigen**vectors**:
Consider a 2x2 matrix:

Way 1:	Way 2: Quicker
Step 1: Write $\mathrm{A}\begin{pmatrix}x\\ y\end{pmatrix}=\lambda\begin{pmatrix}x\\ y\end{pmatrix}$. Step 2: Plug in A and value of λ found Step 3: Do the calculation on the left Step 4: Equate one of the elements (upper or lower)	Step 1: write $\mathrm{A}-\lambda I$ Step 2: Plug in λ and you'll obviously have a matrix Step 3: Multiply this matrix by $\begin{pmatrix}x\\ y\end{pmatrix}$ i.e act on the vector with the matrix and set equal to 0. Solve for x in terms of y and choose any value for x or y

Find the eigenvalues and corresponding eigenvectors of the matrix $A=\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}$

To find eigenvalues:
$\mathrm{A}-\lambda I=\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}-\lambda\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}-\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}=\begin{pmatrix}2-\lambda & 5\\ -1 & -4-\lambda\end{pmatrix}$
$\det(\mathrm{A}-\lambda I)=\begin{vmatrix}2-\lambda & 5\\ -1 & -4-\lambda\end{vmatrix}=(2-\lambda)(-4-\lambda)-5(-1)=\lambda^2+2\lambda-3$
solve $\det(\mathrm{A}-\lambda I)=0$
$\lambda^2+2\lambda-3=0$
$(\lambda-1)(\lambda+3)=0$
$\lambda=1,\lambda=-3$
The eigenvalues of A are 1 and -3

To find eigenvectors:
Write $\mathrm{A}\begin{pmatrix}x\\ y\end{pmatrix}=\lambda\begin{pmatrix}x\\ y\end{pmatrix}$

$\lambda=1$ $\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=1\begin{pmatrix}x\\ y\end{pmatrix}$ $\begin{pmatrix}2x+5y\\ -x-4y\end{pmatrix}=\begin{pmatrix}x\\ y\end{pmatrix}$ Equate the upper elements $2x+5y=x$ $x=-5y$ Let $y=1\Rightarrow x=-5$ So, an eigenvector corresponding to 1 is $\begin{pmatrix}-5\\ 1\end{pmatrix}$ Note: any multiple of this vector is also an eigenvector of A with eigenvalue 1	$\lambda=-3$ $\begin{pmatrix}2 & 5\\ -1 & -4\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=-3\begin{pmatrix}x\\ y\end{pmatrix}$ $\begin{pmatrix}2x+5y\\ -x-4y\end{pmatrix}=\begin{pmatrix}-3x\\ -3y\end{pmatrix}$ Equate the upper elements $2x+5y=-3x$ $5x=-5y$ $x=-y$ Let $y=1\Rightarrow x=-1$ So, an eigenvector corresponding -3 is $\begin{pmatrix}-1\\ 1\end{pmatrix}$ Note: any multiple of this vector is also an eigenvector of A with eigenvalue -3

Alternate method: Get zero on one side and row reduce (this is more useful for when you have larger matrices)

$\begin{pmatrix} 2x+5y-x \\ -x-4y-y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} x+5y \\ -x-5y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} 1 & 5 \\ -1 & -5 \end{pmatrix}$

In RREF

$\begin{pmatrix} 1 & 5 \\ 0 & 0 \end{pmatrix}$

Write circle variables in terms of un-circled like we first learnt for a general solution

$x + 5y = 0$

$x = -5y$

$y = y$

i.e $\begin{pmatrix} x \\ y \end{pmatrix} = y\begin{pmatrix} -5 \\ 1 \end{pmatrix}$

$\begin{pmatrix} -5 \\ 1 \end{pmatrix}$ is the eigenvector

Alternate method: Get zero on one side and row reduce (this is more useful for when you have larger matrices)

$\begin{pmatrix} 2x+5y+3x \\ -x-4y+3y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} 5x+5y \\ -x-y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} 5 & 5 \\ -1 & -1 \end{pmatrix}$

In RREF

$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

Write circle variables in terms of un-circled like we first learnt for a general solution

$5x + 5y = 0$

$x = -y$

$y = y$

i.e $\begin{pmatrix} x \\ y \end{pmatrix} = y\begin{pmatrix} -1 \\ 1 \end{pmatrix}$

$\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is the eigenvector

You are sometimes asked to find a normalised eigenvector. If $\boldsymbol{a} = \begin{pmatrix} a \\ b \end{pmatrix}$ is an eigenvector of a matrix A, then the unit vector

$\hat{\boldsymbol{a}} = \begin{pmatrix} \frac{a}{|a|} \\ \frac{b}{|b|} \end{pmatrix}$ is a normalised eigenvector of A

2.1.3 Eigenspace and Generalised Eigenspace

The eigenspace of (a square matrix) A corresponding to λ is the collection of all vectors $\boldsymbol{x}$ that satisfy $A\,\boldsymbol{x} = \lambda I$, or equivalently, $(A-\lambda I)\boldsymbol{x}=\boldsymbol{0}$. The generalized eigenspace of A corresponding to λ is the collection of all vectors $\boldsymbol{x}$ for which there exists a positive integer k such that $(A-\lambda I)^k\,\boldsymbol{x}=\boldsymbol{0}$. The former is contained in the latter, but need not be equal. For example, with

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

and λ=1, you can check easily that the eigenspace (associated eigenspace) consists only of the vectors of the form $(x,0)$ for some arbitrary x; whereas the generalized eigenspace is the larger collection of all vectors (x,y) with x and y both arbitrary.

We "need" generalized eigenspaces because the eigenspaces in general do not suffice to describe the entire space (the generalized eigenspaces may not suffice either).

So, If A is an $n \times n$ square matrix, with λ an eigenvalue the union of 0 vector and set of all eigenvectors corresponding to eigenvalues λ is a subspace of $\mathbb{R}^n$ known as the eigenspace for λ. To get the eigenspace, find the eigenvectors and write each in terms of the general solution and write as a basis of each eigenvector. Do eigenspace for each eigenvalue. We denote the Eigenspace (associated eigenspace) as:

$$E_\lambda = \{v \in \mathbb{F}^n : \mathbf{A}\mathbf{v} = \lambda\mathbf{v} \text{ i.e } \mathbf{A}\mathbf{v} - \lambda\mathbf{v} = \mathbf{0} \text{ i.e } (\mathbf{A}-\lambda\mathbf{I})\boldsymbol{v} = \mathbf{0}\ \}.$$ Notice how this is just the $\mathrm{Ker}(A-\lambda I)$

We denote the Generalised Eigenspace as:

$$V_n(\lambda) = \{v \in \mathbb{F}^n : (\mathbf{A}-\lambda\mathbf{I})^n\boldsymbol{v} = \mathbf{0}\ \}.$$ Notice how this is just the $\mathrm{Ker}(\mathbf{A}-\lambda\mathbf{V})^n$.

The generalised eigenspace is just $\mathrm{Ker}(\mathbf{A}-\lambda\mathbf{I})^n$, so keep raising $\boldsymbol{A}-\lambda\boldsymbol{I}$ to a power until you get the zero matrix.

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

λ=5, -1 are the eigenvalues (you can check this if you wish)

To find eigenvectors:

Let's do it for $\lambda = 5$

Write $A\begin{pmatrix} x \\ \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix}$

$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 5\begin{pmatrix} x \\ y \end{pmatrix}$

$\begin{pmatrix} x+2y \\ 4x+3y \end{pmatrix} = \begin{pmatrix} 5x \\ 5y \end{pmatrix}$

Equate the upper elements

$x + 2y = 5x$

$4x = 2y$

Let $y = 1 \Rightarrow x = \frac{1}{2}$

So, an eigenvector corresponding to $\lambda = 5$ is $\begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}$

$E_5 = \{v \in \mathbb{F}^n : \boldsymbol{Av} = 5\boldsymbol{v} \text{ i.e } \boldsymbol{Av} - \lambda\boldsymbol{v} = \mathbf{0}\ \}$

$E_5 = \left\{\begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}\right\} = span\left\{\begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}\right\}$

We can multiply the vector by 2 and use integer elements, to make this look neater. Since a span is a linear combination of given vectors, you can safely multiply each vector by the scalars of your choice to 'tidy it up".

$span\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$

Let's do it for $\lambda = -1$

Write $A\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix}$

$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -\begin{pmatrix} x \\ y \end{pmatrix}$

$\begin{pmatrix} x+2y \\ 4x+3y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}$

Equate the upper elements

$x + 2y = -x$

$x = -y$

Let $y = 1 \Rightarrow x = -1$

So, an eigenvector corresponding to $\lambda = -1$ is $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$

$E_{-1} = \left\{\begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\} = span\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$

i.e E_{-1} has basis $\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$

i.e. E_5 has basis $\left\{\begin{pmatrix}1\\2\end{pmatrix}\right\}$	

2.1.4 Minimal Polynomial And Multiplicities

Consider $T: \mathbb{R}^3 \to \mathbb{R}^3$ with some matrix A.
The minimal polynomial $m_T(x)$, is the polynomial with the least degree such that it

- divides the characteristic polynomial $ch_T(x)$
- shares the same roots as the characteristic polynomial (i.e. has the same eigenvalues)
- Is monic (the coefficient of the term of highest degree is equal to 1)
- Takes the least power of the characteristic polynomial to get zero i.e. the least power such that $m_T(A) = 0$

$ch_T(x)$=characteristic polynomial= $(x-3)(x-2)^2$. Find the minimal polynomial $m_T(x)$
$(x-3)(x-2)$ and $(x-3)(x-2)^2$ are the only options
Note: $(x-2), (x-3)$ *and* $(x-2)^2$ are not options since they don't have the same roots. We need to choose at least 1 representation from each factor.

The characteristic polynomial is $(1-x)(x-2)^2$ where matrix A= $\begin{pmatrix}3 & 1 & 1\\ -1 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}$

We know that the
There are 2 possibilities for the minimal polynomial
$m_T(x) = (1-x)(x-2)$ *or* $(1-x)(x-2)^2$
Which of these is correct?
$(1-x)(x-2)$

$$\left[\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} - \begin{pmatrix}3 & 1 & 1\\ -1 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}\right]\left[\begin{pmatrix}3 & 1 & 1\\ -1 & 1 & -1\\ 0 & 0 & 1\end{pmatrix} - \begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}\right]$$

$$\left[\begin{pmatrix}-2 & -1 & -1\\ 1 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}\right]\left[\begin{pmatrix}1 & 1 & 1\\ -1 & -1 & -1\\ 0 & 0 & -1\end{pmatrix}\right] = \begin{pmatrix}1 & -1 & 0\\ -1 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}$$

$(1-x)(x-2)^2$

$$\left[\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} - \begin{pmatrix}3 & 1 & 1\\ -1 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}\right]\left[\begin{pmatrix}3 & 1 & 1\\ -1 & 1 & -1\\ 0 & 0 & 1\end{pmatrix} - \begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}\right]$$

$$\left[\begin{pmatrix}-2 & -1 & -1\\ 1 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}\right]\left[\begin{pmatrix}1 & 1 & 1\\ -1 & -1 & -1\\ 0 & 0 & -1\end{pmatrix}\right]^2$$

$$\left[\begin{pmatrix}-2 & -1 & -1\\ 1 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}\right]\left[\begin{pmatrix}0 & 0 & -1\\ 0 & 0 & 1\\ 0 & 0 & 1\end{pmatrix}\right] = \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

$(1-x)(x-2)^2$ gives 0 so this is the minimal polynomial

It can be long to work out the minimal polynomial when the powers are high e.g $ch_T(x) = (1-x)^4$.
Could compute $1-x, (1-x)^2, (1-x)^3$ *and* $(1-x)^4$. We'd stop at the first one that gave us 0 which might take time.
Instead try to get the matrix $A - \lambda I = A - I$ into upper triangular form with zeros on the diagonal.

So, let's say $A - \lambda I$ was $\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{pmatrix}$

When we raise an upper triangular matrix to a power the numbers get pushed up 1 position each time (as long as start with zeros on the diagonal). So

$(A-\lambda I)^2$ pushes up one place $\begin{pmatrix}0 & 0 & 2 & 0\\ 0 & 0 & 0 & 6\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$

$(A-\lambda I)^3$ pushes up another place $\begin{pmatrix}0 & 0 & 0 & 6\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$

$(A-\lambda I)^4$ gets rid of all numbers $\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$

so $M_T(x) = (1-x)^4$.

2.1.5 Multiplicities

The power of the eigenvalue (number of times the eigenvalue is repeated) is the algebraic multiplicity i.e the powers of the characteristic polynomial for each eigenvalue.

The sum of all powers of the eigenvalues is known as the geometric multiplicity. This is the dimension of the **associated** eigenspace, not the generalised eigenspace. See the eigenspaces section 2.1.3 to understand this.

If for every eigenvalue of A, the geometric multiplicity equals the algebraic multiplicity, then A is said to be diagonalizable. If geometric multiplicity = algebraic multiplicity then diagonalisable.

3 Diagonalisation

Calculations with matrices can often be simplified by reducing a matrix to a given form. In this section you will learn how to reduce some matrices to diagonal form.

The idea of eigenvalues and diagonalization is to approximate general matrices by diagonal matrices, this is to say that matrices are zero everywhere except perhaps on the main diagonal. The eigenvalues are the values that occur on the main diagonal in such an approximation. Granted, not every matrix can be diagonalised (that is to say equal to a diagonal matrix after a basis change), but it can be cast into Jordan form (see section 3.1.3), which also requires the knowledge of eigenvalues.

Any non-zero elements must be on the diagonal! The diagonal can of course have zeros on it

For example, $\begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}, \begin{pmatrix}a & 0\\ 0 & b\end{pmatrix}, \begin{pmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & 0\end{pmatrix}$

Once you've tested for diagonalization, you can just take the eigenvalues and arbitrarily place them in in the i, j, cells to produce a diagonal matrix. The order doesn't matter.

$\begin{pmatrix}1 & -3 & 3\\ 3 & -5 & 3\\ 6 & -6 & 4\end{pmatrix}$ Eigenvalues are $\lambda_1 = 4, \lambda_2 = \lambda_3 = -2$

Diagonal matrix becomes $\begin{pmatrix}-2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 4\end{pmatrix}$ or $\begin{pmatrix}-2 & 0 & 0\\ 0 & 4 & 0\\ 0 & 0 & -2\end{pmatrix}$ or $\begin{pmatrix}4 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -2\end{pmatrix}$

If $Ch_T(x) = (x-1)^4(x-2)^2(x-3)^3$ and $M_T(x) = (x-1)(x-2)(x-3)$

$$\begin{pmatrix}1 &&&&&&&& \\ & 1 &&&&&&& \\ && 1 &&&&&& \\ &&& 1 &&&&& \\ &&&& 2 &&&& \\ &&&&& 2 &&& \\ &&&&&& 3 && \\ &&&&&&& 3 & \\ &&&&&&&& 3\end{pmatrix}$$

We mentioned earlier than the dimension of the matrix will be the power of the characteristic polynomial. Now it should make sense since the diagonals are formed with the eigenvalues.

Once a matrix is in diagonal you can find determinant quickly!

Also once you diagonalise a matrix you can quickly work out powers of matrix. Essentially you just need to raise the entries on the diagonal to the necessary power.

3.1.1 Process Of Diagonalisation

A square matrix which can be reduced to a diagonal form is called a diagonalisable matrix. Not every matrix can be diagonalised, although any $n \times n$ matrix with n distinct eigenvalues can be. Obviously, we can use row/column operations to try and put matrix into diagonal form, but that's long. We want a systematic method to say whether diagonalisable or not. There is another method.

To reduce a given square matrix A into diagonal form we take the following steps:
Step 1; Find the eigenvectors and eigenvalues of A
Step 2: Form a matrix P which consist of the eigenvectors of A **(we form the matrix using the eigenvalues as column vectors)**
Step 3: Find P^{-1}
So, we were able to find a matrix P such that transformed the matrix A into a diagonal matrix i.e. a transformation matrix.
Step 4: A diagonal matrix D is given by $P^{-1}AP$ i.e. D=$P^{-1}AP$. This is equivalent to saying that $A = PDP^{-1}$ (matrix decomposition)
Meaning, if we can find a D such that $P^{-1}AP = D$ where P is the matrix consisting of the eigenvectors of A, the matrix A is diagonalisable.

More formally we say, consider a $n \times n$ matrix A
Step 1: Find the characteristic polynomial of A (solve det(A-$\lambda I) = 0$) to get eigenvalues (i.e. find eigenvalues of the matrix A and their multiplicities from the characteristic polynomial)
Step 2: For each eigenvalue of A, find the basis of the eigenspace E_λ
Step 3: If there is an eigenvalue that gives the geometric multiplicity of λ, dim(E_λ) less than the algebraic multiplicity of λ, then the matrix A is not diagonalizable
Step 4: Combine all basis vectors for all eigenspaces to obtain the linearly independent eigenvectors $v_1, v_2,\ldots, v_n$
Step 5: Define the non-singular matrix S=$[v_1 v_2 \ldots v_n]$
Step 6: Define the diagonal matrix D, where the (i, j)-entry is the eigenvalue, such that the i^{th} column vector v_i is in the eigenspace E_λ. Then the matrix A will finally be diagonalized as P^{-1}AP=D
In other words, any matrix with distinct eigenvalues is diagonalizable (this means V has a basis consisting of eigenvectors of A).

Note: Every upper triangular matrix with distinct elements on the diagonal is diagonalizable.

$$\begin{pmatrix} 1 & 1 & 2 \\ 4 & 2 & -3 \\ 4 & 2 & 3 \end{pmatrix}$$

$$\begin{vmatrix} 1 & 1 & 2 \\ 4 & 2 & -3 \\ 4 & 2 & 3 \end{vmatrix} = (1-\lambda)\left((2-\lambda)(3-\lambda)+6\right) - \left(4(3-\lambda)+12\right) + 2\big(8-4(2-\lambda)\big) = -(\lambda+1)(\lambda-3)(\lambda-4)$$

Eigenvalues $-1,3,4$

$$\begin{pmatrix} 1 & 1 & 2 \\ 4 & 2 & -3 \\ 4 & 2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{pmatrix} x+y+2z \\ 4x+2y-3z \\ 4x+2y+3z \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$\lambda = -1$:

$$\begin{pmatrix} x+y+2z \\ 4x+2y-3z \\ 4x+2y+3z \end{pmatrix} = \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix}$$

Equating the top elements: $x+y+2z = -x \Leftrightarrow 2x+y+2z = 0$
Equating the middle elements: $4x+2y-3z = -y \Leftrightarrow 4x+3y-3z = 0$
Equating the bottom elements: $4x+2y+3z = -z \Leftrightarrow 4x+2y+4z = 0$
Note: only needed a pair, not all three

Pick a pair so we can get a relationship with 2 variables
$4x+2y+4z = 0$
$\underline{4x+3y-3z = 0}$
$-y+7z = 0$

Let $z = 1$:
$-y+7 = 0$
$y = 7$
Sub into either equation: $2x+y+2z = 0$
$2x+7+2 = 0$
$2x = -9$
$x = -\frac{9}{2}$
Corresponding eigenvector is

$$\begin{pmatrix} -\frac{9}{2} \\ 7 \\ 1 \end{pmatrix} = \begin{pmatrix} -9 \\ 14 \\ 2 \end{pmatrix}$$

$\lambda = 3$:

$$\begin{pmatrix} x+y+2z \\ 4x+2y-3z \\ 4x+2y+3z \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \\ 3z \end{pmatrix}$$

Equating the top elements: $x+y+2z = 3x \Leftrightarrow 2x-y-2z = 0$
Equating the middle elements: $4x+2y-3z = 3y \Leftrightarrow 4x-y-3z = 0$
Equating the bottom elements: $4x+2y+3z = 3z \Leftrightarrow 4x+2y = 0$

The last gives
$y = -2x$

Let $x = 1$:
$y = -2$
$z = 2$

Corresponding eigenvector is

$$\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$$

Doing the same for $\lambda = 4$: gives eigenvector $\begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}$

We now build the matrix from all the column vectors

So $P = \begin{pmatrix} -9 & 1 & -1 \\ 14 & -2 & 1 \\ 2 & 2 & -2 \end{pmatrix}$ is the transformation matrix

$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ is the diagonal matrix

Check: $P^{-1}AP = \begin{pmatrix} -\frac{1}{10} & 0 & \frac{1}{20} \\ -\frac{3}{2} & -1 & \frac{1}{4} \\ -\frac{8}{5} & -1 & -\frac{1}{5} \end{pmatrix}\begin{pmatrix} 1 & 1 & 2 \\ 4 & 2 & -3 \\ 4 & 2 & 3 \end{pmatrix}\begin{pmatrix} -9 & 1 & -1 \\ 14 & -2 & 1 \\ 2 & 2 & -2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

Note:

Inside $P = \begin{pmatrix} -9 & 1 & -1 \\ 14 & -2 & 1 \\ 2 & 2 & -2 \end{pmatrix}$ there is a basis of eigenvectors.

In other words, any matrix with distinct eigenvalues is diagonalizable (this means V has a basis consisting of eigenvectors of A).

Note:
As you've already seen, when you reduce a matrix A to a diagonal matrix D, the elements on the diagonal are the eigenvalues of A.
Don't confuse the matrix P which diagonalises a given matrix, A, with the diagonal matrix D. P is formed from the eigenvectors of A, and D has the eigenvalues of A on its leading diagonal. The above process for diagonalising a matrix relies on finding the inverse of P. For larger matrices this can be a time-consuming process. If a matrix is symmetric you can diagonalise it more easily. A matrix, A, is symmetric if $A = A^T$. The elements of a symmetric matrix are symmetric with respect to the leading diagonal. If a matrix A is symmetric you can carry out orthogonal diagonalisation. The procedure for orthogonal diagonalisation of a symmetric matrix is as follows:
Step 1: Find the normalised eigenvectors of A
Step 2: Form a matrix P which consist of the normalised eigenvectors of A
Step 3: Write down P^T
Step 4: A diagonal matrix D is given by P^TAP

3.1.2 How to Check Whether Diagonalisable

Diagonalisable is equivalent to saying

- There exists a basis of eigenvectors.
- The minimal polynomial has distinct roots (distinct factors). This is VERY USEFUL to find out whether a matrix diagonalisation is possible.

In general, any matrix whose eigenvalues are distinct can be diagonalised.

- If there is a repeated eigenvalue, whether or not the matrix can be diagonalised depends on the eigenvectors.
- If there are just two eigenvectors (up to multiplication by a constant), then the matrix cannot be diagonalised.
- If the unique eigenvalue corresponds to an eigenvector **v**, but the repeated eigenvalue corresponds to an entire *plane*, then the matrix can be diagonalised, using **v** together with any two vectors that lie in the plane.
- If all three eigenvalues are repeated, then things are much more straightforward: the matrix can't be diagonalised unless it's already diagonal.

One further nice characterization with the characteristic and minimal polynomial is this. A matrix or linear map is diagonalizable over the field F if and only if its minimal polynomial is a product of distinct linear factors over F.
So first, you can find the characteristic polynomial. If the characteristic polynomial itself is a product of linear factors over F, then you are lucky, no extra work needed, the matrix is diagonalizable. If not, then use the fact that minimal polynomial divides the characteristic polynomial, to find the minimal polynomial. This may not be easy, depending on degree of characteristic polynomial.

3.1.3 Jordan Normal Form

Jordan Normal Form (also called Jordan Canonical Form) is a representation of a linear transformation over a finite-dimensional complex vector space by a particular kind of upper triangular matrix. Every such linear transformation has a unique Jordan canonical form, which has useful properties: it is easy to describe and well-suited for computations. Jordan canonical form can be thought of as a generalization of diagonalizability to arbitrary linear transformations (or matrices); indeed, the Jordan canonical form of a diagonalizable linear transformation (or a diagonalizable matrix) is a diagonal matrix.

Our diagonalisation process mentioned earlier only worked when the square matrix was diagonalisable. How can we generalise this for all square matrices? Instead, we choose a Jordan normal form. Recall earlier $A = PDP^{-1}$ when A diagonalisable, now we can say $A = PJP^{-1}$ where J is a matrix in Jordan Normal Form.

If I choose a square matrix with complex valued entries, then there always exists a Jordan Normal form. It is not uniquely given in general, meaning there could be several Jordan Normal Forms which should be similar in some sense. The complex numbers also include the real numbers so we could have the matrix which only has real numbers as entries. However, this does not necessarily mean we only have real numbers as entries. If we have a Jordan normal form it also means we have an invertible matrix P as before such that we have the matrix decomposition $A = PJP^{-1}$. If A is diagonalisable then Jordan form has to be the diagonal matrix D mentioned previously. The Jordan Normal Form means J could have a diagonalisable matrix, but in general it is not.

A Jordan block is a square matrix of the following form

$$J_\lambda = \begin{pmatrix} \lambda & 1 & & & & & & & \\ & \lambda & 1 & & & & & & \\ & & \lambda & 1 & & & & & \\ & & & \lambda & 1 & & & & \\ & & & & \lambda & 1 & & & \\ & & & & & \lambda & 1 & & \\ & & & & & & \lambda & 1 & \\ & & & & & & & \lambda & 1 \\ & & & & & & & & \lambda \end{pmatrix}$$

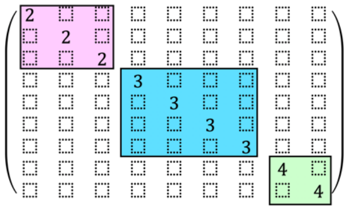
for some complex number λ

Remember the algebraic multiplicity tells how many times each eigenvalue appears (how often we find the eigenvalue as a zero in the characteristic polynomial) and that the sum of algebraic multiplicities is size of matrix.
For diagonalisable matrices we know the eigenvalues appear on the diagonal and the order doesn't matter.
For the eigenvalues are $\lambda_1 = 4, \lambda_2 = \lambda_3 = -2$ i.e $(\lambda - 4)^2$.
Diagonal matrix looks like $\begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ or $\begin{pmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix}$ or $\begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$

What happens if we want to group the same eigenvalues in each block and let's say we want the eigenvalues to increase on the diagonal so the order is fixed. Putting the eigenvalues into these groups is called putting into Jordan blocks, so we have Jordan blocks on each diagonal. It is important to realise that

- each distinct eigenvalue gives us a Jordan block.
- the algebraic multiplicity gives us the size of the corresponding whole Jordan block
- each Jordan box has 1 above the diagonals

Let's say we have the eigenvalues {2,2,2,3,3,3,3,4,4}



So now we know how many Jordan blocks (3) and the sizes of each.

What happens inside each Jordan block? They are independent so doesn't matter which one you start the calculations with!

Recall the geometric multiplicity. It is the size of each Jordan box inside the block. It only tells us the number of the Jordan boxes, NOT the size.
If diagonalisable the algebraic multiplicity = geometric multiplicity

Let's look at the smaller Jordan boxes within each block.

- Look at eigenvalue 2 (3x3 block):
  3 possible geometric multiplicities
  - Geometric multiplicity = algebraic multiplicity = 3 so 3 Jordan boxes (diagonalisable). Remember zeros go outside when diagonalisable.

$$\begin{matrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{matrix}$$

  - Geometric multiplicity 2 so 2 Jordan boxes – 2 in one box and 1 in the other.

$$\begin{matrix} 2 & 1 & \square \\ \square & 2 & \square \\ \square & \square & 2 \end{matrix}$$

  - Geometric multiplicity 1 so 1 Jordan box – 3 in one block.

$$\begin{matrix} 2 & 1 & \square \\ \square & 2 & 1 \\ \square & \square & 2 \end{matrix}$$

- Look at eigenvalue 3 (4x4 block):
  4 possible geometric multiplicities
  - 4 Jordan boxes (diagonalisable). Remember zeros go outside. Geometric multiplicity = algebraic multiplicity =4

$$\begin{matrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{matrix}$$

  - Geometric multiplicity 3 so 3 Jordan boxes – 2 in one box and 1 in the other two boxes. In a Jordan normal block we have 1's above the diagonals.

$$\begin{matrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{matrix}$$

  - Geometric multiplicity 2 so 2 Jordan boxes – 2 in one box and 2 in another OR 3 in one block and 1 in another. Now we have 2 options. Knowing the algebraic and geometric multiplicity doesn't help us to know exactly how the block looks. **This block is too big. We now need more calculations than just the multiplicities!**

$$\begin{matrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{matrix} \qquad \begin{matrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{matrix}$$

  - Geometric multiplicity 1 so 1 Jordan box – 4 in one block.

$$\begin{matrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{matrix}$$

So, we have seen in for the case of $\lambda = 3$, knowing the algebraic and geometric multiplicity doesn't help us to know how the block looks.

We would continue in the same way for $\lambda = 4$

Recall the Generalised Eigenspace as: $V_n(\lambda) = \{v \in \mathbb{F}^n : (\boldsymbol{A} - \lambda \boldsymbol{V})^n = \boldsymbol{0}\ \}$. Notice how this is just the $Ker(\boldsymbol{A} - \lambda \boldsymbol{I})^n$.

Steps to get into JNF:
Step 1: Find all the eigenvalues: solve $det(A - \lambda I) = 0$
Step 2: work out the algebraic and geometric multiplicity. The geometric multiplicity is calculating the kernel: ker $(A - \lambda I)$. Might be done at this point, but in some cases need more that the geometric multiplicity as displayed why in the example above.
Step 3: If needed: calculate ker $(A - \lambda I)^2$, ker $(A - \lambda I)^3$, ... Keep going until matrix no longer changes i.e. becomes the zero matrix. i.e. we need the dimension of the generalised eigenspace aka rank of $(A - \lambda I)$ or rather the nullity which is the dim ker($A - \lambda I$)). Need to put into RREF first like you've already seen.
Repeats steps 2 and 3 for each eigenvalue

$$\begin{pmatrix} 3 & 1 & 0 & 1 \\ -1 & 5 & 4 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Find eigenvalues: $det(A - \lambda I) = 0 \Longrightarrow (2 - \lambda)^4 (4 - \lambda)^3$
$\lambda_1 = 2$ with algebraic multiplicity 1 so Jordan block has size 1x1 and we are done

$\lambda_2 = 4$ with algebraic multiplicity 3 so Jordan block has size 3x3
we already know a lot about the Jordan Normal form for this: the block sizes and that outside the blocks are zeros. We just don't know how many blocks inside the big Jordan block with 4's in it.
Either have
3 Jordan boxes i.e. geometric multiplicity 3
2 Jordan boxes .e. geometric multiplicity 2
1 Jordan box i.e. geometric multiplicity 1
We need to calculate the geometric multiplicity to know which case we're in

$$ker(A - \lambda I) = ker\begin{pmatrix} -1 & 1 & 0 & 1 \\ -1 & 1 & 4 & 1 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = ker\begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = ker\begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = ker\begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We gave 1 pivot (leading one) and 2 free variables so dim(ker)=2=geometric multiplicity
So only one option here for the 3x3 block: 2 boxes and a 2x2 box and 1x1 box

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & -1 & 8 & -3 \\ 0 & 2 & 0 & 7 & 5 \\ 0 & 0 & 2 & 7 & 5 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Find eigenvalues: $det(A - \lambda I) = 0 \Longrightarrow (2 - \lambda)^5$
$\lambda_1 = 2$ with algebraic multiplicity 5 so Jordan block has size 5

Algebraic multiplicity =5

Geometric multiplicity = $\ker(A - \lambda I) = ker\begin{pmatrix} 0 & 1 & -1 & 8 & -3 \\ 0 & 0 & 0 & 7 & 5 \\ 0 & 0 & 0 & 7 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = ker\begin{pmatrix} 0 & 1 & -1 & 8 & -3 \\ 0 & 0 & 0 & 7 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = ker\begin{pmatrix} 0 & 1 & -1 & 8 & -3 \\ 0 & 0 & 0 & 1 & \frac{5}{7} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

IN RREF with 3 free variables (and 2 circled fixed ones) so geometric multiplicity = 3. This means 3 boxes in Jordan block. There are 2 different possibilities for the size of these 3 boxes: block is 5x5 so could have one large and two small OR just two 2x2 boxes and one small box. In order to decide which is correct one for our example, we need to calculate the generalized eigenspace

$$\ker(A-\lambda I)^2 = ker\begin{pmatrix} 0 & 1 & -1 & 8 & -3 \\ 0 & 0 & 0 & 7 & 5 \\ 0 & 0 & 0 & 7 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}^2 = ker\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$\dim(\ker(A-\lambda I)^2) = 5$

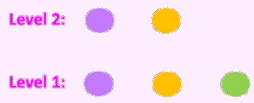
This a good result since algebraic multiplicity = 5 and might have needed to calculate $\ker(A-\lambda I)^3$ *and* $\ker(A-\lambda I)^4$ *and* $\ker(A-\lambda I)^5$ until we reached the algebraic multiplicity of 5 i.e. until we got the zero matrix.

So 2nd level was enough and dimension of generalised eigenspace is 2

Level 1: $\ker(A-\lambda I) = ker\begin{pmatrix} 0 & 1 & -1 & 8 & -3 \\ 0 & 0 & 0 & 1 & \frac{5}{7} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ and 3 free variables so geometric multiplicity =3

Level 2: $\ker(A-\lambda I)^2 = ker\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ and 5 free variables so geometric multiplicity =5

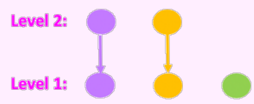
As a diagram:



Level 1: GM = 3 so 3 points

Level 5: G =5 but we only use 2 points since only interested in the new dimensions we get. Already have 3 and to get 5 we add 2 more.

The levels get connected by Jordan chains of different lengths



Lengths 2, 2, and 1. Lengths are the sizes of the Jordan boxes

1 eigenvalue so 1 Jordan block. Inside block have Jordan boxes given by Jordan chains, so (2x2) (2x2) and (1x1)

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Note: we can change the order of the boxes only.

Useful hints/facts:

- Dimension of whole matrix is the sum of the powers of the factors of h_T
- Each distinct eigenvalue gives us a Jordan block
- The algebraic multiplicity gives us the size of the corresponding whole Jordan block
  The power of each factor from $M_T(x)$ tells us the number of Jordan blocks. We know how many eigenvalues in each block
- Each whole Jordan block looks like a $n \times n$ matrix with $\lambda's$ on the diagonal
- Each Jordan block has 1's above the diagonals (0's if diagonalisable, i.e. if same number of boxes as the block size)
- Total number of blocks is $dimv_1(\lambda)$ since each block has the eigenvalue with all zeros below which is the dimension of Ker(A-λI)
- At least one block of size $b \times b$ where this is the largest block. where b = power of min polynomial
- $v_n(\lambda_i)$ says at least a block of size n also.
- Quick way to calculate $dimV_i's$: Calculating Ker$(A-\lambda I)^n$ can take a while
  Ker$(A-\lambda I)^n$ always puts zeros on the diagonal i.e. the kernel sends the diagonals to zero. When we raise an upper triangular matrix to a power the numbers get pushed up 1 position each time (as long as we start with zeros on the diagonal). So, we can use the rule that each power raises the number to get the dim and basis for a kernel

3.1.4 Cayley Hamilton Theorem

The characteristic polynomial expression of a real or complex square matrix will be equal to the zero matrix i.e. a matrix satisfies its own characteristic polynomial. In other words, apply the characteristic polynomial to any matrix and get zero.

So, what does the theorem do for us, besides allowing us to say, "Hey, I know a matrix solution to this polynomial"? Well, it's simple to see that every matrix $n \times n$ has to be a zero of some polynomial of degree at most n^2, simply because the space of $n \times n$ matrices has dimension n^2. The Cayley-Hamilton Theorem says that you can find such a polynomial of much smaller degree. Another way to see this is as follows. For each fixed vector v, the vectors $v, Av, \ldots, A^n v$ cannot be linearly independent and so there is a polynomial p of degree at most n such that P(A)$v = 0$. However, this polynomial depends on

v. The Cayley-Hamilton Theorem gives you a polynomial that works for all vectors v. Finally, for applications, having a polynomial p such that p(A)=0 allows you to compute all powers of A as a linear combination of I, A ,..., A^{n-1}. Indeed assuming p monic, you can write $p(X) = x^n + q(X)$, with q of degree less than n. So $A^n = -q(A)$. Then $A^{n+1} = -Aq(A)$. If A^nappears in $Aq(A)$, replace it by $-q(A)$. Do the same for $A^{n+2} = A \times A^{n+1}$, etc.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix}$$

$$A-\lambda I = \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2 \\ 3 & 7-\lambda \end{pmatrix}$$

$$ch_T = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 \\ 3 & 7-\lambda \end{vmatrix} = (1-\lambda)(7-\lambda) - 2(3) = \lambda^2 - 8\lambda + 1$$

$$ch_T(A) = \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix}^2 - 8\begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix} + 1 = \begin{pmatrix} 7 & 16 \\ 24 & 55 \end{pmatrix} - \begin{pmatrix} 8 & 16 \\ 24 & 56 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

This theorem can also be used to find the inverse of matrices.

If you would like to go a bit further with linear algebra you should look up the following

- Inner, direct and semi-direct products
- Direct sums
- Linear, bilinear, multilinear and quadratic forms
- Adjoints
- Duals

These topics will be covered in the document called Linear Algebra 2.