

Diagonalization

1.1 Eigenvectors

We have seen that when multiplying a matrix with a vector we often get a completely different vector as output.

$$\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} -7 \\ 2 \end{bmatrix}$$

However, every matrix has a collection of vectors which output is particularly nice. These are the so called **eigenvectors** of the matrix. Look for example at

$$\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 6 \end{bmatrix} = 3 \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

The matrix gave us the same vector, scaled by a number, 3.

i Definition 1.1.1

Let A be a $n \times n$ matrix, $\mathbf{x} \in \mathbb{R}^n$ a non-zero vector and $\lambda \in \mathbb{C}$ a scalar. \mathbf{x} is an **eigenvector** of A with **eigenvalue** λ , if

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

! Remark 1.1.1

Note that we need \mathbf{x} to not be the zero vector for it to be an eigenvector. There is nothing special about $A\mathbf{0} = \lambda\mathbf{0}$ since $A\mathbf{0}$ is always $\mathbf{0}$.

Also note that we have that λ can be a complex number. In fact the definition of eigenvectors is also interesting if we let both the vector and matrix to be complex-valued as well, however, we will restrict ourselves to only letting the eigenvalue being possibly complex for now.

Let us observe that if we have an eigenvector \mathbf{x} to a matrix A , then any multiple of \mathbf{x} is still an eigenvector with the same eigenvalue.

$$A(c\mathbf{x}) = c(A\mathbf{x}) = c(\lambda\mathbf{x}) = \lambda(c\mathbf{x})$$

Hence, it is really easy to find $A^n\mathbf{x}$ as long as \mathbf{x} is an eigenvector of A .

$$\begin{aligned} A^n\mathbf{x} &= A^{n-1}(A\mathbf{x}) = A^{n-1}(\lambda\mathbf{x}) \\ &= \lambda(A^{n-1}\mathbf{x}) = \lambda(A^{n-1}(A\mathbf{x})) = \lambda(A^{n-2}(\lambda\mathbf{x})) \\ &= \lambda^2(A^{n-2}\mathbf{x}) = \dots \\ &\dots = \lambda^n\mathbf{x} \end{aligned}$$

📊 Example 1.1.1

Let A be the matrix

$$\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}$$

and \mathbf{v} be the eigenvector

$$\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

with eigenvalue 3. Find $A^{50}\mathbf{v}$.

$$A^{50}\mathbf{v} = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}^{50} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 3^{50} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

This is in fact the property we are most interested in here in this course. Let us look at how we can use this fact to find $A^n\mathbf{y}$ for any \mathbf{y} as long as A has a sufficient amount of linearly independent eigenvectors.

Example 1.1.2

The matrix A given by

$$\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}$$

has eigenvector \mathbf{v} of eigenvalue 3 given by

$$\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

and eigenvector \mathbf{u} of eigenvalue -5 given by

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

\mathbf{u} and \mathbf{v} are linearly independent (check) and any vector \mathbf{x} can be written as a linear combination of them. Find $A^5\mathbf{x}$ for

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

We first want to write \mathbf{x} as a linear combination of \mathbf{u} and \mathbf{v} , that is, we want to find a and b such that

$$a \begin{bmatrix} -1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

to find a and b , we reduce the augmented matrix

$$\left[\begin{array}{cc|c} -1 & 3 & 1 \\ 2 & 2 & 0 \end{array} \right] \sim \left[\begin{array}{cc|c} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \end{array} \right]$$

and get $a = -1/4$ and $b = 1/4$. Now,

$$\begin{aligned} \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}^5 \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}^5 \left(-\frac{1}{4} \begin{bmatrix} -1 \\ 2 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right) \\ &= -\frac{1}{4} \left(\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}^5 \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right) + \frac{1}{4} \left(\begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}^5 \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right) \\ &= -\frac{1}{4} \cdot (-5)^5 \begin{bmatrix} -1 \\ 2 \end{bmatrix} + \frac{1}{4} \cdot 3^5 \begin{bmatrix} 3 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} -599 \\ 1684 \end{bmatrix} \end{aligned}$$

Finding eigenvectors

Now, how do we find eigenvectors? Let us look at the defining property:

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

In this equation we can move $\lambda\mathbf{x}$ to the left hand side and get

$$A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0},$$

or equivalently

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

where I is the identity matrix. So we are looking for non-zero solutions to $(A - \lambda)\mathbf{x} = \mathbf{0}$. By Theorem ??, we know that this is equivalent to $\det(A - \lambda I) = 0$. When writing out $\rho_A(\lambda) = \det(A - \lambda)$, we get a polynomial of order n with variable λ , which we call the **characteristic polynomial of A** .

We know therefore that the eigenvalues of A are given as the roots of $\rho_A(\lambda)$. To find the eigenvectors associated to these, we need to solve $(A - \lambda I)\mathbf{x} = \mathbf{0}$ for all the eigenvalues λ , or equivalently determining the the null space $\text{Null}(A - \lambda I)$.

Theorem 1.1.1

Let A be a $n \times n$ -matrix.

- (a) The eigenvalues of A are the solutions λ of the equation

$$\rho_A(\lambda) = \det(A - \lambda I_n) = 0.$$

- (b) If λ is an eigenvalue of A , then the associated eigenvectors are given as the

non-trivial solutions of the equation

$$(A - \lambda I_n) \cdot \mathbf{x} = \mathbf{0}.$$

Example 1.1.3

Let us now use theorem 1.1.1 to find the eigenvalues and eigenvectors to the matrix

$$A = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}$$

The eigenvalues are found by solving

$$\det(A - \lambda I_2) = 0,$$

Let us first see how the matrix $A - \lambda I_2$ looks like:

$$A - \lambda I_2 = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1 - \lambda & 3 \\ 4 & -3 - \lambda \end{bmatrix}$$

The characteristic polynomial is

$$\begin{aligned} \rho_A(\lambda) = \det(A - \lambda I_2) &= \begin{vmatrix} 1 - \lambda & 3 \\ 4 & -3 - \lambda \end{vmatrix} \\ &= (1 - \lambda)(-3 - \lambda) - 3 \cdot 4 \\ &= \lambda^2 + 2\lambda - 15. \end{aligned}$$

This means that we can solve the second order equation

$$\lambda^2 + 2\lambda - 15 = 0$$

to find the eigenvalues. We solve it as usual and get:

$$\lambda = \frac{-2 \pm \sqrt{2^2 - 4 \cdot (-15)}}{2} = -1 \pm 4$$

We therefore have the eigenvalues: 3 and -5.

We find the eigenvectors associated to 3 by solving the equation $(A - 3I_2)\mathbf{x} = \mathbf{0}$. We solve it by reducing the matrix $(A - 3I_2)$:

$$A - 3I_2 = \begin{bmatrix} -2 & 3 \\ 4 & -6 \end{bmatrix} \sim \begin{bmatrix} -2 & 3 \\ 0 & 0 \end{bmatrix}$$

We get a free variable and the solution is

$$\mathbf{x} = t \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

for all numbers t . The eigenvectors corresponding to 3 are therefore all vectors in

$$\text{span} \left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\},$$

except the zero vector.

We find all eigenvectors to -5 by solving the equation $(A + 5I_2)\mathbf{x} = \mathbf{0}$. We can solve this equation by reducing the matrix $(A + 5I_2)$:

$$A + 5I_2 = \begin{bmatrix} 6 & 3 \\ 4 & 2 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

We get a free variable, and the solution is

$$\mathbf{x} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \cdot t$$

for all numbers t . The eigenvectors corresponding to the eigenvalue -5 are therefore all the vectors in

$$\text{span} \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\},$$

except the zero vector.

i Definition 1.1.2

A **diagonal matrix** is a quadratic matrix where all numbers outside of the diagonal are 0, that is a matrix on the form:

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \vdots & a_{nn} \end{bmatrix}$$

The eigenvalues of a diagonal matrix are easily obtained

Theorem 1.1.2

The eigenvalues of a diagonal matrix are the numbers along the diagonal.

Proof. The characteristic polynomial of the diagonal matrix A is

$$\det(A - \lambda \cdot I_n) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda).$$

So the numbers a_{11}, \dots, a_{nn} are the solutions of the equation $\det(A - \lambda \cdot I_n) = 0$. \square

We can also note that the eigenvectors gives yet another condition for the invertability of a matrix

Theorem 1.1.3

A $n \times n$ -matrix A has 0 as an eigenvalue if and only if it is not invertible.

We can add this to our running theorem ?? and thus have

Theorem 1.1.4

Let A be a $n \times n$ -matrix. The following statements are equivalent

1. A is invertible
2. $\det A \neq 0$
3. A do not have 0 as an eigenvalue
4. The columns of A are linearly independent
5. The rows of A are linearly independent
6. $\text{rank}A = n$
7. $A\mathbf{x} = \mathbf{b}$ has a unique solution ($\mathbf{x} = A^{-1}\mathbf{b}$)
8. $A\mathbf{x} = \mathbf{b}$ has a solution for every \mathbf{c} .
9. The reduced row echelon form of A is I_n
10. A has a pivot element in every row
11. $A\mathbf{x} = \mathbf{0}$ has only one solution, $\mathbf{x} = \mathbf{0}$
12. $\text{Col}(A) = \mathbb{R}^n$

13. $\text{Row}(A) = \mathbb{R}^n$

A last useful result we can note is that the eigenvectors corresponding to different eigenvalues are linearly independent:

 **Theorem 1.1.5**

Let A be a $n \times n$ matrix. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t$ be eigenvectors of A associated to different eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_t$. Then the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t$ are linearly independent.

 **Example 1.1.4**

We find the eigenvalues of

$$A = \begin{bmatrix} -8 & 0 & 6 \\ 12 & 4 & -6 \\ -20 & 0 & 14 \end{bmatrix},$$

and the corresponding eigenvectors.

The characteristic polynomial of A is:

$$\begin{aligned} \det(A - \lambda I_3) &= \begin{vmatrix} -8 - \lambda & 0 & 6 \\ 12 & 4 - \lambda & -6 \\ -20 & 0 & 14 - \lambda \end{vmatrix} \\ &= (4 - \lambda) \cdot \begin{vmatrix} -8 - \lambda & 6 \\ -20 & 14 - \lambda \end{vmatrix} \\ &= (4 - \lambda) \left((-8 - \lambda)(14 - \lambda) + 6 \cdot 20 \right) \\ &= (4 - \lambda)(\lambda^2 - 6\lambda + 8) \end{aligned}$$

We find the eigenvalues of A by solving the third degree equation

$$(4 - \lambda)(\lambda^2 - 6\lambda + 8) = 0.$$

This equation is equivalent to

$$4 - \lambda = 0 \quad \text{or} \quad \lambda^2 - 6\lambda + 8 = 0.$$

The second degree equation $\lambda^2 - 6\lambda + 8 = 0$ have solutions

$$\lambda = \frac{6 \pm \sqrt{6^2 - 4 \cdot 8}}{2} = 3 \pm 1,$$

so we have two eigenvalues: 2 og 4.

We find the eigenvectors by solving the equations

$$(A - 2I_3)\mathbf{x} = \mathbf{0} \quad \text{og} \quad (A - 4I_3)\mathbf{x} = \mathbf{0}.$$

For $\lambda = 2$ we reduce the matrix $A - 2I_3$:

$$\begin{aligned} \begin{bmatrix} -8-2 & 0 & 6 \\ 12 & 4-2 & -6 \\ -20 & 0 & 14-2 \end{bmatrix} &\sim \begin{bmatrix} -10 & 0 & 6 \\ 12 & 2 & -6 \\ -20 & 0 & 12 \end{bmatrix} \\ &\sim \begin{bmatrix} -10 & 0 & 6 \\ 12 & 2 & -6 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} -5 & 0 & 3 \\ 6 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

The first row gives the equation $-5x_1 + 3x_3 = 0$. This is satisfied for example when $x_1 = 3$ and $x_3 = 5$. These two values gives in the second equation that $x_2 = -3$. The result is that the eigenvectors of 2 are the non-zero vectors in

$$\text{span} \left\{ \begin{bmatrix} 3 \\ -3 \\ 5 \end{bmatrix} \right\}.$$

To find the eigenvectors of 4 we reduce $A - 4I_3$:

$$\begin{aligned} \begin{bmatrix} -8-4 & 0 & 6 \\ 12 & 4-4 & -6 \\ -20 & 0 & 14-4 \end{bmatrix} &\sim \begin{bmatrix} -12 & 0 & 6 \\ 12 & 0 & -6 \\ -20 & 0 & 10 \end{bmatrix} \\ &\sim \begin{bmatrix} -2 & 0 & 1 \\ 2 & 0 & -1 \\ -2 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} -2 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Here we have that both x_2 and x_3 are free variables. If we choose $x_2 = 1$ and $x_3 = 0$, we get $x_1 = 0$. If we choose $x_3 = 2$, we have from the first row that

$x_1 = 1$. In this case x_2 can be chosen freely, so we can choose $x_2 = 0$. This gives us that the eigenvectors of 4 are all the non-zero vectors of

$$\text{span} \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}.$$

⚠ Remark 1.1.2

In an effort to declutter our arguments we often use the notion of **eigenspaces** of a matrix. If λ is an eigenvalue of A , then the eigenspace, E_λ , of λ is exactly the null space we have used to find the eigenvectors:

$$E_\lambda = \text{Null}(A - \lambda I)$$

The dimension of an eigenspace E_λ is called the **geometric multiplicity** of λ .

The power of $(\lambda - \lambda_i)$ in the characteristic polynomial $\rho_A(\lambda)$ of a matrix A , is called the **algebraic multiplicity** of λ , i.e. if

$$\rho_A(\lambda) = (\lambda - 1)^2(\lambda - 4)^3(\lambda + 1)$$

then the eigenvalue 1 has algebraic multiplicity 2, the eigenvalue 4 has algebraic multiplicity 3 and the eigenvalue -1 has algebraic multiplicity 1.

1.2 Diagonalization

Consider the diagonal matrix

$$D = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix}.$$

It is easy to multiply D with vectors in \mathbb{R}^2 or \mathbb{C}^2 . But, it is also easy to multiply D with itself for example D^5 can be found as

$$D^5 = D \cdot D \cdot D \cdot D \cdot D = \begin{bmatrix} 3^5 & 0 \\ 0 & (-5)^5 \end{bmatrix} = \begin{bmatrix} 243 & 0 \\ 0 & -3125 \end{bmatrix}.$$

If we try the same with A , that is to find A^5 for

$$A = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix},$$

the calculations are more cumbersome.

Now, recall that we found that A have eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -5$ with eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

This means that we have

$$A\mathbf{v}_1 = 3\mathbf{v}_1 \quad \text{og} \quad A\mathbf{v}_2 = (-5)\mathbf{v}_2.$$

Rewriting these two equations by having the eigenvalues as the diagonal in a matrix D and the eigenvectors as the columns in a matrix P we get:

$$D = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \quad \text{og} \quad P = \begin{bmatrix} 3 & 1 \\ 2 & -2 \end{bmatrix},$$

such that

$$AP = PD.$$

Observe that P is invertible with inverse

$$P^{-1} = \frac{1}{8} \begin{bmatrix} 2 & 1 \\ 2 & -3 \end{bmatrix}.$$

We can therefore multiply both sides of $AP = PD$ with P^{-1} and get

$$A = PDP^{-1}.$$

Now we can calculate A^k for all k with the formula

$$\begin{aligned} A^k &= \overbrace{(PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1})}^{n \text{ times}} \\ &= PD \overbrace{(P^{-1}P)}{=I_2} D \overbrace{(P^{-1}P)}{=I_2} \cdots D \overbrace{(P^{-1}P)}{=I_2} DP^{-1} \\ &= PD^k P^{-1} \end{aligned}$$

i Definition 1.2.1

A $n \times n$ -matrix A is **diagonalizable** if there exists a diagonal matrix D and an invertible matrix P such that

$$A = PDP^{-1}.$$

We say that P diagonalize A in these cases.

Not all matrices are diagonalizable. We need a method to check whether we can diagonalize A .

 **Theorem 1.2.1**

A $n \times n$ -matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

Proof. We only show that enough eigenvectors implies diagonalizability. Assume A has n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ and corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. For every eigenvector we have

$$A\mathbf{v}_k = \lambda_k \mathbf{v}_k.$$

As above, we can organize these n equations into a matrix equation

$$AP = PD,$$

where

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

and

$$P = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n].$$

Since $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are linearly independent, the $n \times n$ -matrix P is invertible. We can therefore find an inverse P^{-1} and we have

$$A = PDP^{-1}.$$

We conclude that A is diagonalizable. □

 **Remark 1.2.1**

Generally, if a $n \times n$ -matrix has n different eigenvalues, then it is diagonalizable since each eigenvalue has at least one eigenvector and we know that eigenvectors belonging to different eigenvalues are linearly independent.

 **Example 1.2.1**

We have already looked at the matrix

$$A = \begin{bmatrix} -8 & 0 & 6 \\ 12 & 4 & -6 \\ -20 & 0 & 14 \end{bmatrix},$$

and know that the eigenvalues to A are 2 and 4. The eigenvectors of 2 are the non-zero vectors in

$$\text{span} \left\{ \begin{bmatrix} 3 \\ -3 \\ 5 \end{bmatrix} \right\}.$$

The eigenvectors of 4 are the non-zero vectors in

$$\text{span} \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}.$$

We observe that we have three linearly independent eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ -3 \\ 5 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ og } \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

This means that A is diagonalizable with the diagonal matrix

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

and invertable matrix

$$P = \begin{bmatrix} 3 & 0 & 1 \\ -3 & 1 & 0 \\ 5 & 0 & 2 \end{bmatrix}.$$

Example 1.2.2

The matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

have eigenvalues

$$\lambda = \pm i$$

with the nonzero vectors in $\text{span} \left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}$ as eigenvectors to i , and the nonzero vectors in $\text{span} \left\{ \begin{bmatrix} 1 \\ i \end{bmatrix} \right\}$ as eigenvectors to $-i$.

This means that A is diagonalizable as a *complex* matrix with diagonal matrix

$$D = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

and invertible matrix

$$P = \begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}.$$

Example 1.2.3

We consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The eigenvalue of A are 1:

$$\det \left(\begin{bmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{bmatrix} \right) = (1 - \lambda)^2.$$

The eigenvectors of 1 are the non-zero vectors of the zero space to the matrix

$$A - I_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

This is the vectors spanned by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We can't construct two linearly independent vectors from only one, so in particular we see that A do not have 2 linearly independent eigenvectors and is therefore not diagonalizable.

We can also formulate our findings on diagonalizability in terms of the algebraic and geometric multiplicity of the eigenvalues:

Theorem 1.2.2

Let A be a $n \times n$ -matrix. If the algebraic multiplicity and geometric multiplicity of each eigenvalue of A is equal, then A is diagonalizable.

Proof. Let

$$\rho_A(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0$$

be the characteristic polynomial of A . It has degree n since A is a $n \times n$ -matrix. Now, by the fundamental theorem of algebra we can decompose $\rho_A(\lambda)$ as follows

$$\rho_A(\lambda) = A(\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2} \cdots (\lambda - \lambda_t)^{\alpha_t}$$

where A is some constant, $\lambda_1, \dots, \lambda_t \in \mathbb{C}$ are eigenvalues (potentially complex) of A , and $\alpha_1 + \alpha_2 + \cdots + \alpha_t = n$.

Now, since the geometric multiplicity of λ_i equals the algebraic multiplicity α_i , we can pick a set of α_i linearly independent eigenvectors belonging to λ_i . Since we know that eigenvectors belonging to different eigenvalues are linearly independent, we can collect the sets of linearly independent eigenvectors of each eigenvalue into a big set of linearly independent eigenvectors.

Further, since $\alpha_1 + \alpha_2 + \cdots + \alpha_t = n$, we have then found n linearly independent eigenvectors of A , and we conclude that A is diagonalizable. \square

Symmetric matrices

A particularly nice family of matrices when it comes to diagonalizability is the real valued symmetric ones. We will see in a bit that they are always diagonalizable.

Definition 1.2.2

A real valued matrix is called **symmetric** if $A = A^T$.

Example 1.2.4

The matrix

$$\begin{bmatrix} 1 & -5 & 7 \\ -5 & 2 & -13 \\ 7 & -13 & 3 \end{bmatrix}$$

is symmetric.

Theorem 1.2.3

Let A be a symmetric $n \times n$ -matrix. Then A has n real valued eigenvalues and A is diagonalizable.

1.3 Jordan form (optional)

⚠ Remark 1.3.1

The readability of this section is not tried optimized, venture on with caution.

What do we do if there aren't enough linearly independent eigenvectors to diagonalize a matrix A ? In this case we would still like to construct an invertible matrix P such that

$$A = PJP^{-1}$$

for a matrix J as diagonal as possible. Such a matrix, it turns out, is a **Jordan form** matrix and is found through the construction of additional **generalized eigenvectors** which is put alongside our original eigenvectors into P .

A matrix of Jordan form can be seen as a big $n \times n$ matrix J subdivided into smaller $l_j \times l_j$ matrices J_j along the diagonal, and zeros elsewhere,

$$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_t \end{bmatrix}$$

the jordan blocks are on the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 \\ & \lambda_i & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \vdots \\ & & & \lambda_i & 1 & 0 \\ & & & & \lambda_i & 1 \\ & & & & & \lambda_1 \end{bmatrix}$$

with only zeros below the diagonal.

Suppose A has t linearly independent eigenvectors. Then it is can be written as PJP^{-1} for a Jordan form matrix with t blocks J_1, \dots, J_t . Each block has an eigenvalue on the diagonal with 1's in the entry right above. The matrix P consists of the generalized eigenvectors of A .

ⓘ Definition 1.3.1

Let A be a $n \times n$ -matrix and λ an eigenvalue of A . A non-zero vector \mathbf{x} is a **generalized eigenvector** of A if for some integer p , we have

$$(A - \lambda I)^p \mathbf{x} = \mathbf{0}$$

we say that \mathbf{x} is a generalized eigenvector of **rank** p if also

$$(A - \lambda I)^{p-1} \mathbf{x} = \mathbf{0}$$

In particular, we have that any eigenvector of λ is a generalized eigenvector of rank 1.

Given a generalized eigenvector \mathbf{v} of rank r belonging to λ , we define the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ iteratively as

$$\begin{aligned} \mathbf{v}_r &= (A - \lambda I)^0 \mathbf{v} = I\mathbf{v} = \mathbf{v} \\ \mathbf{v}_{r-1} &= (A - \lambda I)\mathbf{v}_r = (A - \lambda I)^1 \mathbf{v} \\ \mathbf{v}_{r-2} &= (A - \lambda I)\mathbf{v}_{r-1} = (A - \lambda I)^2 \mathbf{v} \\ &\vdots \\ \mathbf{v}_2 &= (A - \lambda I)\mathbf{v}_3 = (A - \lambda I)^{r-2} \mathbf{v} \\ \mathbf{v}_1 &= (A - \lambda I)\mathbf{v}_2 = (A - \lambda I)^{r-1} \mathbf{v} \end{aligned}$$

Note that \mathbf{v}_1 is an eigenvector as $\mathbf{v}_1 \neq \mathbf{0}$ and $(A - \lambda I)\mathbf{v}_1 = (A - \lambda I)^r \mathbf{v} = \mathbf{0}$. The vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ form a **chain of generalized eigenvectors of length** r .

Theorem 1.3.1

The vectors in a chain of generalized eigenvectors are linearly independent.

Theorem 1.3.2

For an eigenvalue λ of algebraic multiplicity k , there exist p chains of generalized eigenvectors,

$$\begin{array}{ccccccc} \mathbf{v}_1^1 & , & \mathbf{v}_1^2 & , & \cdots & , & \mathbf{v}_1^p \\ \mathbf{v}_2^1 & & \mathbf{v}_2^2 & & & & \mathbf{v}_2^p \\ \vdots & & \vdots & & & & \vdots \\ \vdots & & \mathbf{v}_{r_2}^2 & & & & \mathbf{v}_{r_p}^p \\ \mathbf{v}_{r_1}^1 & & & & & & \end{array}$$

such that the collection of all these vectors

$$\{\mathbf{v}_1^1, \dots, \mathbf{v}_{r_1}^1, \mathbf{v}_1^2, \dots, \mathbf{v}_{r_2}^2, \dots, \mathbf{v}_1^p, \dots, \mathbf{v}_{r_p}^p\}$$

forms a linearly independent set, and $\sum_i^p r_i = k$. Here r_i denotes the length of the i th chain.

sth chain of eigenvalue λ_i :

$$\begin{array}{c} \leftarrow r_s \rightarrow \\ \uparrow r_s \\ \left[\begin{array}{cccccc} \lambda_i & 1 & 0 & 0 & \cdots & 0 \\ & \lambda_i & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \vdots \\ & & & \lambda_i & 1 & 0 \\ & & & & \lambda_i & 1 \\ & & & & & \lambda_i \end{array} \right] \end{array}$$

Example 1.3.1

Let

$$A = \begin{bmatrix} 0 & 1 \\ -4 & 4 \end{bmatrix}$$

Then the characteristic polynomial of A is $\rho_A(\lambda) = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$. The eigenspace

$$E_2 = \text{Null} \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix} = \text{Null} \begin{bmatrix} 1 & -1/2 \\ 0 & 0 \end{bmatrix} = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

is only one-dimensional, i.e. the geometric multiplicity of 2 is 1 which is lower than the algebraic multiplicity which is 2. The matrix is therefore not diagonalizable.

We observe however that $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is in the column space of $A - 2I$, i.e. the equation

$$\begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

has a solution, for example $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Further, since

$$(A - 2I_2)\mathbf{v}_2 = \mathbf{v}_1$$

we see that

$$A\mathbf{v}_2 = \mathbf{v}_1 + 2\mathbf{v}_2 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

combining this with

$$A\mathbf{v}_1 = 2\mathbf{v}_1 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

we have

$$A \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

If we denote $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$ by P and observe this matrix is invertible, we have

$$A = P \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} P^{-1}$$

and observe that we have decomposed A into its Jordan form.