Definition 6.29. Eigenspace

The solution set of the LES $(A - \lambda\mathbb{1})\mathbf{x} = \mathbf{o}$*, which means* $\operatorname{Ker}(A - \lambda\mathbb{1})$*, is called the* *eigenspace* *with respect to the eigenvalue* λ *and denoted by* $\operatorname{Eig}(\lambda)$*. Each* ***nonzero*** *vector* $\mathbf{x} \in \operatorname{Eig}(\lambda) \setminus \{\mathbf{o}\}$ *is an eigenvector w.r.t. the eigenvalue* λ*.*

VL16 ↓

$\in \mathbb{C}^{2\times 2}$ $\quad p_A(\lambda) = (3-\lambda)(2-\lambda) - 2 = 4 - \underbrace{5}_{\operatorname{tr}(A)}\lambda + \lambda^2$

$= (\lambda - 1)(\lambda - 4)$

Example 6.30. Consider $A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$: $\mathbf{x}_i \neq \mathbf{o}$ is an eigenvalue for λ_i with $i \in \{1, 2\}$ if

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \qquad \text{i.e.} \qquad (A - \lambda_i \mathbb{1})\mathbf{x}_i = \mathbf{o}.$$

$\Rightarrow \lambda_1 = 4, \ \lambda_2 = 1$

Hence, we have to solve the LES $(A - \lambda_1\mathbb{1})\mathbf{x}_1 = \mathbf{o}$ and $(A - \lambda_2\mathbb{1})\mathbf{x}_2 = \mathbf{o}$.

$$\underline{\lambda_1 = 4}: \qquad A - \lambda_1 \mathbb{1} = \begin{pmatrix} 3 - \lambda_1 & 2 \\ 1 & 2 - \lambda_1 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix},$$

$$(A - \lambda_1\mathbb{1})\mathbf{x}_1 = \mathbf{o}$$

"complex" span

$$\Leftrightarrow \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix} \mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \mathbf{x}_1 \in \operatorname{Span}\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right)$$

In the same manner:

$$\underline{\lambda_2 = 1}: \qquad A - \lambda_2 \mathbb{1} = \begin{pmatrix} 3 - \lambda_2 & 2 \\ 1 & 2 - \lambda_2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix},$$

$$(A - \lambda_2\mathbb{1})\mathbf{x}_2 = \mathbf{o}$$

$$\Leftrightarrow \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \mathbf{x}_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \mathbf{x}_2 \in \operatorname{Span}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$$

Definition 6.31. Multiplicities

Let $A \in \mathbb{C}^{n\times n}$ *be square matrix. Then the characteristic polynomial can be written as:*

$$p_A(z) = (\lambda_1 - z)^{\alpha_1} \cdot (\lambda_2 - z)^{\alpha_2} \cdots (\lambda_k - z)^{\alpha_k} \tag{6.2}$$

where $\lambda_1, \ldots, \lambda_k$ *are pairwise different. The natural number* α_j *above is called:*

$$\alpha(\lambda_j) := \alpha_j \quad \textit{algebraic multiplicity of } \lambda_j$$

and tells you how often the eigenvalue λ_j *occurs in the characteristic polynomial. We also define*

$$\gamma(\lambda_j) := \dim(\operatorname{Eig}(\lambda_j)) = \dim(\operatorname{Ker}(A - \lambda_j\mathbb{1})) \quad \textit{geometric multiplicity of } \lambda_j$$

Remark: Recipe for calculating eigenvectors

Let $A \in \mathbb{C}^{n\times n}$ be a square matrix.

(1) The eigenvalues λ are the zeros of the characteristic polynomial p_A of A. In other words, the solutions of

$$p_A(\lambda) = \det(A - \lambda \mathbb{1}) = 0.$$

(2) If A is real, then $p_A(\lambda)$ is a real polynomial. If it has a complex zero $\lambda \notin \mathbb{R}$, then its conjugate $\overline{\lambda}$ is also a zero,

(3) If one eigenvalue is found, we can reduce the characteristic polynomial by equating coefficients (or polynomial division).

(4) The eigenvectors $\mathbf{x}$ are given by the solutions of the LES $(A - \lambda\mathbb{1})\mathbf{x} = \mathbf{o}$ for each eigenvalue, where only the nonzero solutions $\mathbf{x} \neq \mathbf{o}$ are interesting.

Example 6.32.

$$p(\lambda) = -\lambda^3 + 5\lambda^2 - 8\lambda + 6$$

- $n = 3$ is odd: "$-\lambda^3$"
- Try some values and find: $\lambda_1 = 3$.

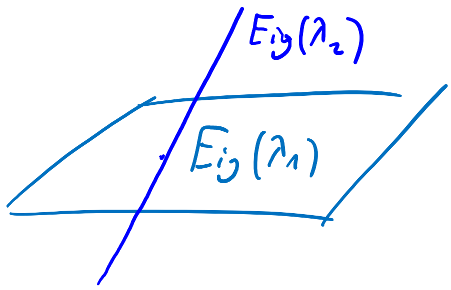
$$(-\lambda^3 + 5\lambda^2 - 8\lambda + 6) : (\lambda - 3) = -\lambda^2 + 2\lambda - 2$$
$$-(-\lambda^3 + 3\lambda^2)$$
$$0 + 2\lambda^2 - 8\lambda$$
$$-(2\lambda^2 - 6\lambda)$$
$$0 - 2\lambda + 6$$
$$-(-2\lambda + 6)$$
$$0$$

$$\Rightarrow p(\lambda) = \underbrace{(-\lambda^2 + 2\lambda - 2)}_{\text{find the zeros now!}}(\lambda - 3)$$

Exercise 6.33.

Let A be a square matrix and λ_1, λ_2 two different eigenvalues. Show that

$$\mathrm{Eig}(\lambda_1) \cap \mathrm{Eig}(\lambda_2) = \{\mathbf{o}\}$$



6.6 The spectral mapping theorem

Let $\lambda \in \mathbb{C}$ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$ corresponding to the eigenvector $\mathbf{x} \in \mathbb{C}^n$, which means $A\mathbf{x} = \lambda\mathbf{x}$. Then we get for the powers:

$$A^2 x = A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda\cdot\lambda\cdot x = \lambda^2\cdot x$$

$$A^3 x = A(A^2 x) = A(\lambda^2 x) = \lambda^2(Ax) = \lambda^2\cdot\lambda x = \lambda^3 x$$

$$\vdots$$

$$A^m x = \lambda^m x \quad \text{for all } m \in \mathbb{N}.$$

We conclude that A^m has also the eigenvector $\mathbf{x}$ but now it corresponds to the eigenvalue λ^m instead of λ.

$$(A^2 + A^3)x = A^2 x + A^3 x = (\lambda^2 + \lambda^3)x$$

Now we could also bring in the addition of the matrices A^0, A^1, A^2, and so on, and get a similar result.

Proposition 6.34. Polynomial spectral mapping theorem

Let $p(\lambda) = p_m\lambda^m + p_{m-1}\lambda^{m-1} + \ldots + p_1\lambda + p_0$ be a polynomial and $A \in \mathbb{C}^{n\times n}$ a square matrix. Putting the matrix A into p (formally), we get the following matrix:

$$p(A) := p_m A^m + p_{m-1}A^{m-1} + \ldots + p_1 A + p_0\mathbb{1} \in \mathbb{C}^{n\times n}$$

It is again an $n \times n$ matrix, and we get

$$\operatorname{spec}\big(p(A)\big) = \{p(\lambda) : \lambda \in \operatorname{spec}(A)\}.$$

Moreover, each eigenvector of A is also an eigenvector of $p(A)$.

Proof: Skript

Example 6.35. Let $A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$. We want to know the eigenvalues of the following matrix $B = 3A^3 - 7A^2 + A - 2\mathbb{1}$.

$\rightsquigarrow$ $\operatorname{spec}(A) = \{4, 1\}$, $x_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $x_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\hookrightarrow$ Spectrum of B:

$$\mu_1 = 3(4)^3 - 7\cdot 4^2 + 4 - 2 = 82$$

$$\mu_2 = 3\cdot 1^3 - 7\cdot 1^2 + 1 - 2 = -5$$

$\rightsquigarrow$ Eigenvector for $\mu_1 = 82$ is $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$

" " $\mu_2 = -5$ is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

A invertible: $\quad Ax = \lambda x$

$\Rightarrow A^{-1}(Ax) = A^{-1}(\lambda x)$

$\Rightarrow x = \lambda \cdot (A^{-1}x) \quad \Rightarrow \boxed{A^{-1}x = \frac{1}{\lambda} \cdot x}$

ev-equation for A^{-1}

$$Ax = \lambda x,\ x \neq 0 \quad \Rightarrow \quad A^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}. \tag{6.3}$$

Rule of thumb:

A^{-1} has the same eigenvector $\mathbf{x}$ as A – but for the eigenvalue λ^{-1} instead of λ.

We simply get:

$$\text{spec}(A^{-1}) = \{\lambda^{-1} : \lambda \in \text{spec}(A)\}.$$

Of course, λ^{-1} is always well-defined since $\lambda \neq 0$. $\quad$ (A inv. $\Leftrightarrow 0 \notin \text{spec}(A)$)

Example 6.36. Let $A = \left(\begin{smallmatrix} 3 & 2 \\ 1 & 2 \end{smallmatrix}\right)$.

$\hookrightarrow A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightsquigarrow A^{-1} = \frac{1}{\det(A)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

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

This matrix has the eigenvalues $\mu_1 = 1/4$ and $\mu_2 = 1$ and the eigenvectors $\mathbf{x}_1 = \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$ and $\mathbf{x}_2 = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$.

Alternatively: $\quad \text{spec}(A) = \{4, 1\}$

$\text{spec}(A^{-1}) = \{\frac{1}{4}, \frac{1}{1}\}$

We do not have to stop here. We can multiply A^{-1} again from the left to equation (6.3) and, doing this repeatedly, we get

$$A^{-2}\mathbf{x} = \lambda^{-2}\mathbf{x}, \qquad A^{-3}\mathbf{x} = \lambda^{-3}\mathbf{x}, \qquad \text{etc.},$$

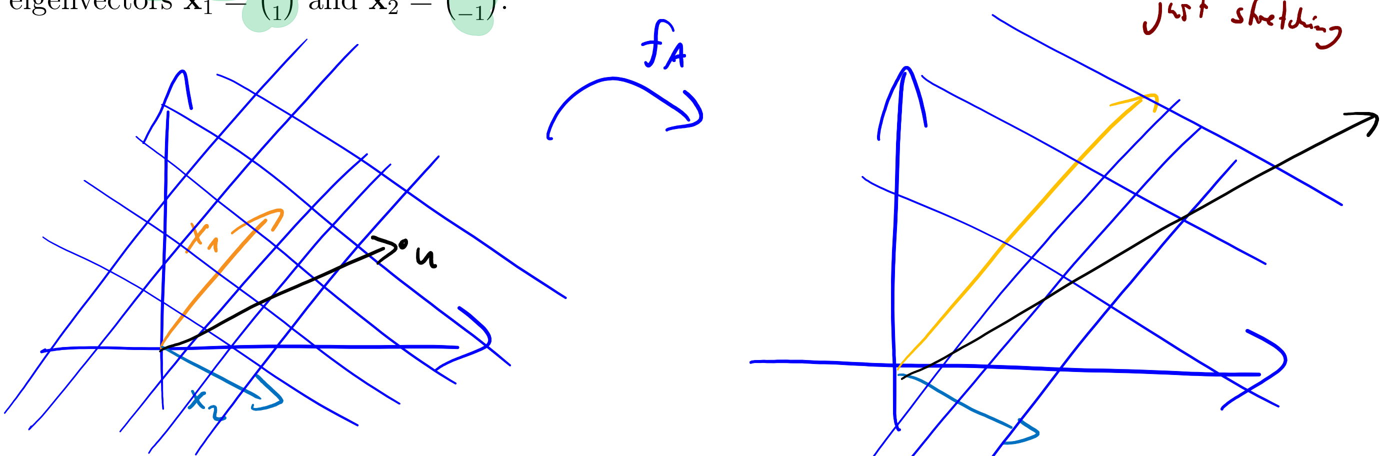
where A^{-n} means $(A^{-1})^n$.

$$Ax = \lambda x,\ x \neq 0 \Rightarrow \quad A^m\mathbf{x} = \lambda^m\mathbf{x} \quad \textit{for all } m \in \mathbb{Z}.$$

Of course, if we can also expand it to linear combinations $\ldots, A^{-2}, A^{-1}, A^0, A^1, A^2, \ldots$ which shows that our spectral mapping theorem is only a special case of a more general one.

6.7 Diagonalisation – the optimal coordinates

We startet this chapter with a two-dimensional picture. Now, we again revisit the 2×2-example $A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$. We know that $\lambda_1 = 4$ and $\lambda_2 = 1$ are the eigenvalues with associated eigenvectors $\mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\mathbf{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.



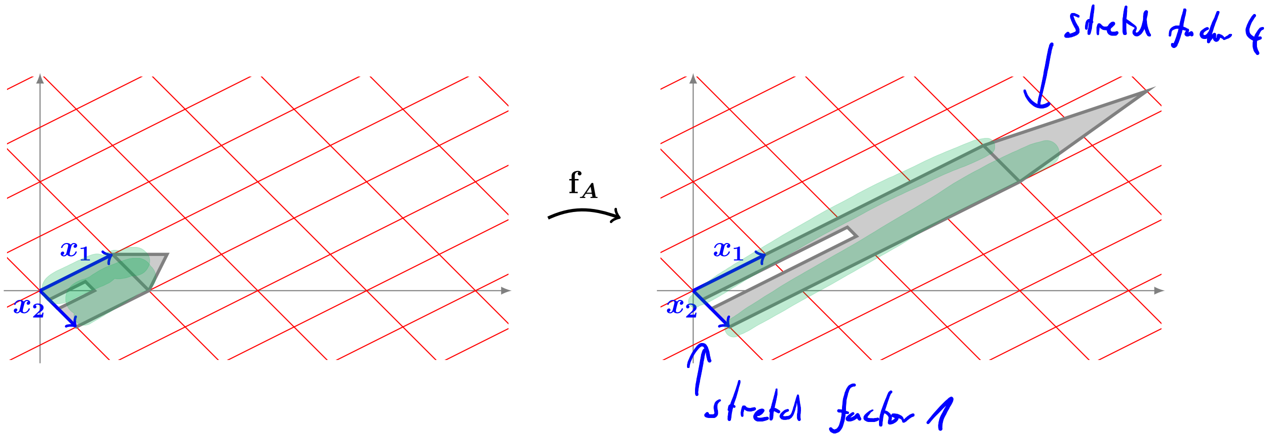
Optimal coordinates for A

By using for $\mathbf{u} \in \mathbb{R}^2$ the linear combination $\mathbf{u} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$ with coefficients $\alpha_1, \alpha_2 \in \mathbb{R}$, we get

$$A\mathbf{u} = A(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1(A\mathbf{x}_1) + \alpha_2(A\mathbf{x}_2) = \alpha_1(4\mathbf{x}_1) + \alpha_2(1\mathbf{x}_2) = 4\alpha_1\mathbf{x}_1 + 1\alpha_2\mathbf{x}_2.$$

The component in $\mathbf{x}_1$-direction, which is α_1, is scaled by the factor $\lambda_1 = 4$, and the $\mathbf{x}_2$-component α_2 is scaled by the factor $\lambda_2 = 1$.

$$A^{100} u = 4^{100} \alpha_1 x_1 + 1^{100} \alpha_2 x_2$$



Generalise this: $A \in \mathbb{C}^{n \times n}$

$\rightsquigarrow \lambda_1, \ldots, \lambda_n \in \mathbb{C}$ (counted with alg. multiplicities)

$\rightsquigarrow x_1, \ldots, x_n \in \mathbb{C}^n$ corresponding eigenvectors

n eigenvalue equations

$$A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1, \quad \ldots, \quad A\mathbf{x}_n = \lambda_n \mathbf{x}_n . \tag{6.4}$$

This is what we can put together into a matrix equation:

$$\underbrace{A\begin{pmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{pmatrix}}_{=:X} = \begin{pmatrix} | & & | \\ A\mathbf{x}_1 & \cdots & A\mathbf{x}_n \\ | & & | \end{pmatrix}$$

$$\overset{(6.4)}{=} \begin{pmatrix} | & & | \\ \lambda_1\mathbf{x}_1 & \cdots & \lambda_n\mathbf{x}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{pmatrix} \underbrace{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}}_{=:D},$$

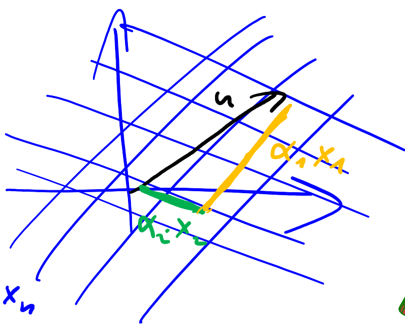
or in short: $AX = XD$. This means that A is similar to a diagonal matrix if X is invertible.

X invertible $\Rightarrow X^{-1}AX = D$

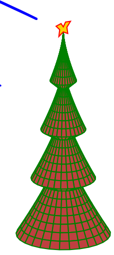
Choose $v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{C}^n$: $(AX)v = (XD)v = X\begin{pmatrix} \lambda_1\alpha_1 \\ \vdots \\ \lambda_n\alpha_n \end{pmatrix}$

$\quad \| $

$A(\underbrace{\alpha_1 x_1 + \cdots + \alpha_n x_n}_{u}) = \lambda_1 \cdot \alpha_1 \cdot x_1 + \cdots + \lambda_n \cdot \alpha_n \cdot x_n$



$u = \alpha_1 x_1 + \alpha_2 x_2$



A "acts as" $\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} \mapsto \begin{pmatrix}\lambda_1\alpha_1\\ \vdots \\ \lambda_n\alpha_n\end{pmatrix}$

$(e_1, \dots, e_n) \xrightarrow{\text{switching}} (x_1, \dots, x_n)$ basis of eigenvectors

Diagonalisation of A

Choose $X = \begin{pmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{pmatrix}$ *and* $D = \begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$. *Then:*

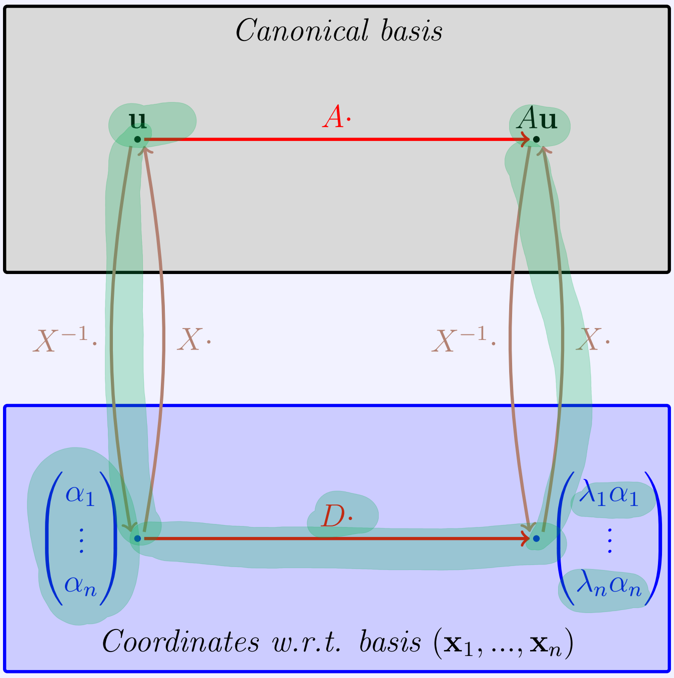
$$AX = XD. \tag{6.5}$$

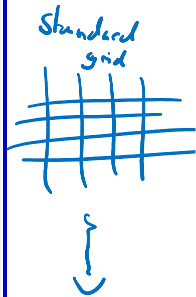
Multiplication (6.5)$\cdot X^{-1}$ *gives:*

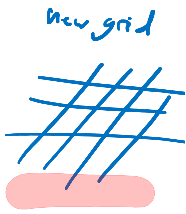
$$A = XDX^{-1} \tag{6.6}$$

and in the same ways $X^{-1}\cdot$(6.5) *gives:*

$$X^{-1}AX = D.$$







$$A^{99} = (XDX^{-1})^{99} = XD\underbrace{X^{-1}X}_{\mathbb{1}}D\underbrace{X^{-1}X}_{\mathbb{1}}D\underbrace{X^{-1}X}_{\mathbb{1}}DX^{-1}\cdots XDX^{-1}$$

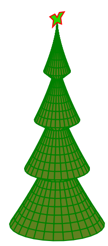
$$= XD^{99}X^{-1}$$

The important question "Is that even possible?" is equivalent to the following:

- Can we write all $\mathbf{u} \in \mathbb{C}^n$ as $\alpha_1\mathbf{x}_1 + \ldots + \alpha_n\mathbf{x}_n$?
- $\mathrm{Span}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \mathbb{C}^n$?
- Is $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ a basis of $\mathbb{C}^n$?
- Is X invertible?

Definition 6.37. Diagonalisability

A square matrix $A \in \mathbb{C}^{n\times n}$ is called diagonalisable *if one can find n different eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{C}^n$ that form a basis of $\mathbb{C}^n$.*



Example 6.38. (a) The matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ has $\mathbf{e}_1$ and $\mathbf{e}_2$ as eigenvectors and they form a basis of $\mathbb{C}^2$. Hence, A is diagonalisable.

(b) The matrix $B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ has $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as eigenvectors and they form a basis of $\mathbb{C}^2$. Hence, B is diagonalisable.

(c) The matrix $C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has only eigenvectors in direction $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and they cannot form a basis of $\mathbb{C}^2$. Hence, C is not diagonalisable.

Choosing a basis consisting of eigenvectors, we know that A acts like a diagonal matrix.

Proposition 6.39. Different eigenvalues $\Rightarrow$ linearly ind. eigenvectors

If $\lambda_1, \ldots, \lambda_k$ are k different eigenvalues of A, then each family $(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ of corresponding eigenvectors is linearly independent.

Recall: $Eig(\lambda_1) \cap Eig(\lambda_2) = \{0\}$

$\Rightarrow (x_1, x_2)$ lin. indep.

Proof:

Induction: over k   base case

Corollary 6.40. n different eigenvalues $\Rightarrow$ diagonalisable

If $A \in \mathbb{C}^{n \times n}$ has n different eigenvalues, then A is a diagonalisable.

Proof. A linearly independent family of n eigenvectors forms a basis for $\mathbb{C}^n$. $\square$

Example 6.41. (a) $A = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$ has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 1$. Corollary 6.40 tells us that A is diagonalisable. We also verify this by looking at the eigenvectors $\mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\mathbf{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, which form a basis of $\mathbb{C}^2$. Hence, $\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} = A = XDX^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}^{-1}$.

(b) The 90°-rotation $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has eigenvalues $\lambda_{1,2} = \pm \mathrm{i}$. From $A - \lambda_{1,2}\mathbb{1} = \begin{pmatrix} \mp\mathrm{i} & -1 \\ 1 & \mp\mathrm{i} \end{pmatrix}$ we conclude the eigenvectors $\mathbf{x}_1 = \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix}$ and $\mathbf{x}_2 = \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix}$, which span $\mathbb{C}^2$.

$$\operatorname{Ker}(A - i\mathbb{1}) = \operatorname{Ker}\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \overset{II - iI}{=} \operatorname{Ker}\begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix} = \operatorname{Span}\left(\begin{pmatrix} i \\ 1 \end{pmatrix}\right) = \operatorname{Span}\left(\begin{pmatrix} -1 \\ i \end{pmatrix}\right)$$

$$\operatorname{Ker}(A + i\mathbb{1}) = \ldots$$

$$X = \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}, \quad X^{-1} = \frac{1}{2}\begin{pmatrix} i & -1 \\ -1 & i \end{pmatrix}$$

$$A = XDX^{-1}, \quad D = \begin{pmatrix} i & \\ & -i \end{pmatrix}$$

$$A^{100}$$

Hence, $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = A = XDX^{-1} = \begin{pmatrix} \mathrm{i} & 1 \\ 1 & \mathrm{i} \end{pmatrix}\begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}\begin{pmatrix} \mathrm{i} & 1 \\ 1 & \mathrm{i} \end{pmatrix}^{-1}$. X and D are strictly complex, while A is a real matrix.

(c) Look at the 3×3 matrices:

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 1 & 6 & 3 \\ -2 & -4 & -2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 8 & 8 & 4 \\ -1 & 2 & 1 \\ -2 & -4 & -2 \end{pmatrix}.$$

If you calculate the characteristic polynomials, you find

$\alpha(4) = 2$

$$p_A(\lambda) = -\lambda^3 + 8\lambda^2 - 16\lambda = -\lambda(\lambda - 4)^2 = p_B(\lambda)$$

$\alpha(0) = 1$

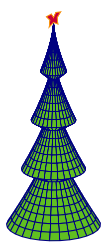
and, hence, the same eigenvalues $\lambda_1 = 0$, $\lambda_2 = 4$ and $\lambda_3 = 4$.

For A, the eigenspaces are:

$$\operatorname{Ker}(A - \lambda_1\mathbb{1}) = \operatorname{Span}\left(\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}\right) \quad \text{and} \quad \operatorname{Ker}(A - \lambda_2\mathbb{1}) = \operatorname{Span}\left(\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}\right),$$

$\gamma(4) = 2$ for A

However for B, the eigenspaces are



$, \gamma(4) = 1$

$$\operatorname{Ker}(B - \lambda_1 \mathbb{1}) = \operatorname{Span}\left(\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}\right) \quad \text{and} \quad \operatorname{Ker}(B - \lambda_2 \mathbb{1}) = \operatorname{Span}\left(\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}\right).$$

While A has three different directions for eigenvectors and is diagonalisable, the matrix B has for $\lambda_{2,3} = 4$ only one direction for eigenvectors. There are too few vectors for a basis and B is not diagonalisable.

(d) Let $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

$\hookrightarrow$ not diagonalisable! $\underline{D = \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}}$

$\underline{A} = X D X^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ↯

Reminder: Algebraic and geometric multiplicity

For each eigenvalue λ of A we consider

- *the algebraic multiplicity of λ, denoted by $\alpha(\lambda)$, given by the multiplicity of λ as zero of p_A, and*
- *the geometric multiplicity of λ, denoted by $\gamma(\lambda)$, given by the dimension of the eigenspace* $\operatorname{Ker}(A - \lambda \mathbb{1})$.

For A from Example 6.41 (c), we find $\alpha(0) = 1 = \gamma(0)$, $\alpha(4) = 2 = \gamma(4)$.
For B from Example 6.41 (c), we get $\alpha(0) = 1 = \gamma(0)$, $\alpha(4) = 2 \neq 1 = \gamma(4)$.

Proposition 6.42. Algebraic vs. geometric multiplicity

Let $A \in \mathbb{C}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ be all eigenvalues of A (not counted with multiplicities). Then:

(a) $\alpha(\lambda_1) + \ldots + \alpha(\lambda_k) = n$.

(b) For all $i = 1, \ldots, k$, we have $1 \leq \gamma(\lambda_i) \leq \alpha(\lambda_i)$.

Therefore, the following claims are equivalent:

(a) A is diagonalisable,

(b) $\gamma(\lambda_1) + \ldots + \gamma(\lambda_k) = n$,

(c) $\gamma(\lambda_i) = \alpha(\lambda_i)$ for all $i = 1, \ldots, k$.

Proof. Exercise. □

Checking diagonalisability $\Rightarrow$ Checking $\gamma(\lambda_i) = \alpha(\lambda_i)$ for all i

$A = A^*$ selfadjoint

For symmetric or selfadjoint matrices, we can improve Proposition 6.39 even more:

Proposition 6.43. $A{=}A^*$: **orthogonal eigenvectors**

Let $A \in \mathbb{C}^{n\times n}$ be selfadjoint, which means $A = A^*$, and let $\lambda, \lambda' \in \mathbb{C}$ be two different eigenvalues of A with corresponding eigenvectors $\mathbf{x}$ and $\mathbf{x}'$, respectively. Then $\mathbf{x} \perp \mathbf{x}'$. (w.r.t. the standard inner product in $\mathbb{C}^n$)

Proof. Since $\langle \mathbf{x}, \lambda'\mathbf{x}'\rangle \overset{(S4)}{=} \overline{\langle \lambda'\mathbf{x}', \mathbf{x}\rangle} = \overline{\lambda'\langle \mathbf{x}', \mathbf{x}\rangle} = \overline{\lambda'}\,\overline{\langle \mathbf{x}', \mathbf{x}\rangle} \overset{(S4)}{=} \overline{\lambda'}\,\langle \mathbf{x}, \mathbf{x}'\rangle$, we have

$$\lambda\langle \mathbf{x}, \mathbf{x}'\rangle = \langle \lambda\mathbf{x}, \mathbf{x}'\rangle = \langle A\mathbf{x}, \mathbf{x}'\rangle \overset{A=A^*}{=} \langle \mathbf{x}, A\mathbf{x}'\rangle = \langle \mathbf{x}, \lambda'\mathbf{x}'\rangle \overset{\text{see above}}{=} \overline{\lambda'}\langle \mathbf{x}, \mathbf{x}'\rangle \overset{\text{Prop. 6.24}}{=} \lambda'\langle \mathbf{x}, \mathbf{x}'\rangle$$

and, hence, $(\lambda - \lambda')\langle \mathbf{x}, \mathbf{x}'\rangle = 0$. This means that the second factor has to be zero. □

Proposition 6.44. $A{=}A^*$: **diagonalisable - ONB of eigenvectors**

Let $A \in \mathbb{C}^{n\times n}$ be selfadjoint, which means $A = A^*$. Then A is diagonalisable, where there is an ONB $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ for $\mathbb{C}^n$ consisting of eigenvectors of A. The matrix

$x_j^* x_i = \langle x_i, x_j \rangle = \delta_{ij}$

$$X = \begin{pmatrix} | & & | \\ \mathbf{x}_1 & \ldots & \mathbf{x}_n \\ | & & | \end{pmatrix}$$

is unitary, i.e. $X^{-1} = X^*$. Therefore, we have:

$$A = XDX^{-1} = XDX^* \qquad \text{and} \qquad D = X^{-1}AX = X^*AX. \tag{6.7}$$

in $\mathbb{C}^n$

Sketch of proof. Use Proposition 6.43 and Gram-Schmidt for each eigenspace to find an ONB of $\mathbb{C}^n$. Then $X^*X = \mathbb{1}$ and also $X^* = X^{-1}$. □

Actually, we could generalise the Proposition from above and equation (6.7). It holds if and only if the matrix A is normal (i.e. $AA^* = A^*A$).

$$A = X \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} X^*$$

Proposition 6.45.

For a diagonalisable $A \in \mathbb{C}^{n \times n}$*, let* $\lambda_1, \ldots, \lambda_n$ *be the eigenvalues counted with algebraic multiplicities. Then*

$$\det(A) = \prod_{i=1}^{n} \lambda_i \quad \text{and} \quad \operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i \,,$$

where $\operatorname{tr}(A) := \sum_{j=1}^{n} a_{jj}$ *is the sum of the diagonal, the so-called trace of* A*.*

Proof. Exercise! □

Remark:

Later, we will see that the result of Proposition 6.45 actually holds for all matrices $A \in \mathbb{C}^{n \times n}$*.*

6.8 Some applications

Here, we look at some of very many possible applications.

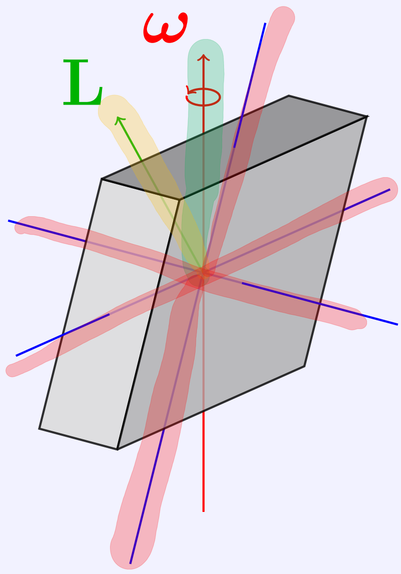
Rotation of boxes

A box of the size $10\text{cm} \times 20\text{cm} \times 30\text{cm}$ *rotates around a axis given by the vector* $\boldsymbol{\omega} \in \mathbb{R}^3$*. The whole box has a angular momentum* $\mathbf{L} \in \mathbb{R}^3$*.*

$\mathbf{L}$ *is given by a linear equation using* $\boldsymbol{\omega}$*, which means*

$$\mathbf{L} = \mathbf{J}\boldsymbol{\omega}$$

with a symmetric matrix $\mathbf{J} \in \mathbb{R}^{3 \times 3}$*, which is called the inertia tensor of the box. The rotation "wobbles" if* $\mathbf{L}$*, which means* $\mathbf{J}\boldsymbol{\omega}$*, is not parallel to the rotation axis* $\boldsymbol{\omega}$*. Of course, we have three special rotation axes given by the eigenvectors of* $\mathbf{J}$*. They are called the principal axes of the box.*



diagonalisation ⟺ find eigenvectors

⟺ find eigenvalues

Curves and areas

Which points $\mathbf{x} = \binom{x}{y} \in \mathbb{R}^2$ *satisfy the equation* $3x^2 + 2\sqrt{3}xy + y^2 + x - \sqrt{3}y = 2$?
Solution: Rewrite the equation as a vector-matrix equation

$$2 = 3x^2 + 2\sqrt{3}xy + y^2 + x - \sqrt{3}y = \underbrace{(x\ y)}_{\mathbf{x}^T} \underbrace{\begin{pmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}}_{=:A\,(=A^T)} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\mathbf{x}} + \underbrace{(1\ \ -\sqrt{3})}_{=:\mathbf{b}^T} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\mathbf{x}}$$

and diagonalise the symmetric matrix A: $\lambda_1 = 4,\ \lambda_2 = 0,\ \mathbf{x}_1 = \frac{1}{2}\binom{\sqrt{3}}{1},\ \mathbf{x}_2 = \frac{1}{2}\binom{-1}{\sqrt{3}}$

$$\begin{pmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} = A = XDX^* = XDX^T = \underbrace{\frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} 4 & \\ & 0 \end{pmatrix}}_{D} \underbrace{\frac{1}{2}\begin{pmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{pmatrix}}_{X^T}.$$

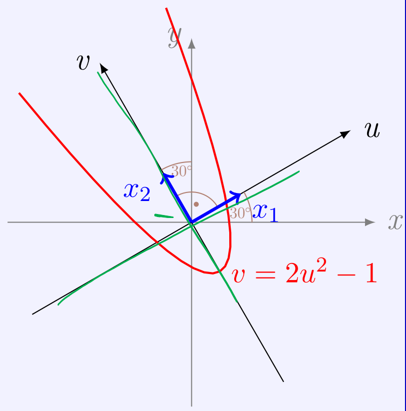
Then we get $2 = \mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} = \mathbf{x}^T(XDX^T)\mathbf{x} + \mathbf{b}^T\mathbf{x} = (\mathbf{x}^T X)D(X^T\mathbf{x}) + \mathbf{b}^T X(X^T\mathbf{x})$.
Setting $\binom{u}{v} = \mathbf{u} := X^T\mathbf{x} = \frac{1}{2}\begin{pmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{pmatrix}\binom{x}{y}$ *simplifies the equation to*

$$2 = \mathbf{u}^T D\mathbf{u} + \mathbf{b}^T X\mathbf{u} = (u\ v)\begin{pmatrix} 4 & \\ & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} + \underbrace{(0\ \ -2)}_{\mathbf{b}^T X}\begin{pmatrix} u \\ v \end{pmatrix} = 4u^2 - 2v.$$

The more complicated equation from above looks at lot simpler in the "optimal" $(\mathbf{x}_1, \mathbf{x}_2)$*-coordinate system:*

$$2 = 4u^2 - 2v, \qquad \textit{also} \qquad v = 2u^2 - 1,$$

There you immediately see that it is a ***parabola****. The transformation we did,* $\binom{x}{y} = \mathbf{x} \mapsto \binom{u}{v} = \mathbf{u} = X^T\mathbf{x}$, *was just a rotation by* 30°.



A simple criterion for definiteness

$n = 2$: $\det(A) = \lambda_1 \lambda_2$.

- $\det(A) > 0 \Rightarrow$ eigenvalues have the same sign $\Rightarrow A$ (pos. or neg.) definite. If $a_{11} = \mathbf{e}_1^T A \mathbf{e}_1 > 0$, then pos., otherwise neg. definite.
- $\det(A) < 0 \Rightarrow A$ indefinite

In general: A symmetric A is positive definite if all *left upper subdeterminants* are positive.

Summary

- All matrices A we considered here were square matrices.
- A vector $\mathbf{x} \neq \mathbf{o}$, which A only scales, which means $A\mathbf{x} = \lambda \mathbf{x}$, is called an *eigenvector*; the corresponding scaling factor λ is called an *eigenvalue*. The set of all eigenvalues of A is called the *spectrum*.
- λ is an eigenvalue of A if and only if $(A - \lambda \mathbb{1})\mathbf{x} = \mathbf{o}$ has non-trivial solutions $\mathbf{x} \neq \mathbf{o}$ (namely the eigenvectors). This is fulfilled if and only if $\det(A - \lambda \mathbb{1}) = 0$.
- For $A \in \mathbb{C}^{n \times n}$, we define $p_A(\lambda) := \det(A - \lambda \mathbb{1})$, the *characteristic polynomial of A*, which is a polynomial of degree n in the variable λ. It has exactly n complex zeros: the eigenvalues of A.
- The eigenvalues λ are in general complex numbers, also the eigenvectors are complex $\mathbf{x} \in \mathbb{C}^n$. All matrices should be considered as $A \in \mathbb{C}^{n \times n}$.
- Also in $\mathbb{C}^n$, we can define inner products. Here, we only use the standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle$, defined by $x_1 \overline{y_1} + \cdots + x_n \overline{y_n}$. Hence we get a new operation for matrices: $A^* := \overline{A^T} = (\overline{a_{ji}})$. It satisfies $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^* \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y}$.
- Checking eigenvalues: Product of all eigenvalues of A is equal to $\det(A)$; the sum is equal to $\operatorname{tr}(A)$.
- The matrix A is invertible if and only if all eigenvalues are nonzero.
- The eigenvalues of a triangular matrix are the diagonal entries. The eigenvalues of a block matrix in triangular form are given by the eigenvalues of the blocks on the diagonal.
- The eigenvalues of A^m are given by the eigenvalues of A to the power of m where $m \in \mathbb{Z}$. The eigenvectors stay the same as for A. For example, $3A^{17} - 2A^3 + 5A^{-6}$ has the eigenvalues $3\lambda^{17} - 2\lambda^3 + 5\lambda^{-6}$, where λ goes through all eigenvalues of A.

The eigenvectors still stay the same.

- A is called *diagonalisable* if it can be written as XDX^{-1}, where D is a diagonal matrix consisting of eigenvalues of A and X gets eigenvectors in the columns. This only works if there are enough eigenvectors directions such that X is invertible.
- A is diagonalisable if and only if for all eigenvalues λ the *algebraic multiplicity* of λ is the same as the *geometric multiplicity*.
- If $A = A^*$, then A is diagonalisable and the eigenvalues are real and eigenvectors can be chosen to be orthonormal.

Happy New Year!