Linear Algebra 2

Hans Sterk (dutch version),
chapter 3, appendix B and english translation: Aart Blokhuis

Faculty of Mathematics and Computer Science
Eindhoven University of Technology

2017-2019

Syllabus for Linear Algebra 2 (2WF30)

Contents

Chapter 1

Linear maps

1.1 Linear maps

1.1.1 In this section we introduce maps between vector spaces that respect the vector space structure (see B.2.1), so called *linear maps*. Such maps are used to study connections between vector spaces. Topics in this section are:

- the notion of a linear map;
- the composition or product, sum, scalar multiple, and inverse of linear maps;
- the null space $\mathcal{N}$ and the range $\mathcal{R}$ of a linear map, and the relation of these and injectivity and surjectivity (see A.1.2);
- to specify a linear map on a given basis (see B.2.5) and
- (very important) the dimension theorem.

Of extreme importance is the connection between matrices (see B.3.1) and linear maps, this will be the subject of the next sections.

Let V and W be vector spaces. A map $\mathcal{A} : V \to W$ associates to *each* vector $\underline{v}$ in V exactly one vector $\mathcal{A}(\underline{v})$ (or $\mathcal{A}\underline{v}$) in W.

1.1.2 Definition. (Linear map) A map $\mathcal{A} : V \to W$ is called *linear* if for all vectors $\underline{x}, \underline{y} \in V$ and all (real) numbers (scalars) λ the following holds ($\underline{0}, +, \lambda$):

$$\begin{aligned}
&\mathcal{A}(\underline{0}) = \underline{0} \ , \\
&\mathcal{A}(\underline{x} + \underline{y}) = \mathcal{A}\underline{x} + \mathcal{A}\underline{y} \ , \\
&\mathcal{A}(\lambda\underline{x}) = \lambda\mathcal{A}\underline{x} \ .
\end{aligned}$$

Equivalently: for all $\underline{x}, \underline{y} \in V$ and all numbers α, β one has

$$\mathcal{A}(\alpha\underline{x} + \beta\underline{y}) = \alpha\mathcal{A}\underline{x} + \beta\mathcal{A}\underline{y} \ .$$

A bijective linear map is also called an *isomorphism*.

1.1.3 Repeatedly applying the definition we see that the image of a linear combination is the same linear combination of the image vectors: $\mathcal{A} : V \to W$ is linear if and only if

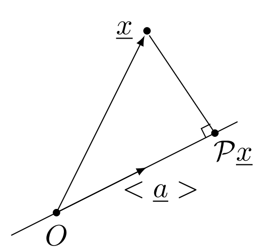
$$\mathcal{A}\left(\sum_{i=1}^{n} \alpha_i \underline{x}_i\right) = \sum_{i=1}^{n} \alpha_i \mathcal{A}\underline{x} \ ,$$

for all vectors $\underline{x}_1, \ldots, \underline{x}_n$ in V and all numbers $\alpha_1, \ldots, \alpha_n$.

Linear maps occur extremely often in practice, even though not always immediately recognized as such. Multiplication with a fixed matrix, orthogonal projection on a subspace of an inner product space, rotation around an axis through the origin in E^3, reflection in a plane through the origin and differentiation in a suitable vector space of functions are all linear. In chapter 2 rotations and reflections will be treated extensively.

1.1.4 **Example. (Orthogonal projection)** Let $l =< \underline{a} >$ be a line (through the origin) in the (real) inner product space V (see B.2.6). The map that associates to each vector in V the orthogonal projection on l we will call $\mathcal{P}$. If $\underline{a}$ has length one, then $\mathcal{P}$ is given by the formula:

$$\mathcal{P}\underline{x} = (\underline{x}, \underline{a})\underline{a} \ ,$$



We verify, using this formula, that this map is linear. Let $\underline{x}, \underline{y} \in V$ and let $\alpha, \beta \in \mathbb{R}$ be arbitrary, then

$$\begin{aligned} \mathcal{P}(\alpha\underline{x} + \beta\underline{y}) &= (\alpha\underline{x} + \beta\underline{y}, \underline{a})\underline{a} &&= \big(\alpha(\underline{x}, \underline{a}) + \beta(\underline{y}, \underline{a})\big)\,\underline{a} \\ &= \alpha(\underline{x}, \underline{a})\underline{a} + \beta(\underline{y}, \underline{a})\underline{a} &&= \alpha\mathcal{P}\underline{x} + \beta\mathcal{P}\underline{y} \ , \end{aligned}$$

showing that the map is linear.

In a similar way one can see that orthogonal projection on a linear subspace of a real inner product space produces a linear map . (If $\{\underline{a}_1, \dots, \underline{a}_k\}$ is an orthonormal basis of the subspace W, then we have the following formula for the orthogonal projection $\mathcal{P}\underline{x}$ of a vector $\underline{x}$:

$$\mathcal{P}\underline{x} = (\underline{x}, \underline{a}_1)\underline{a}_1 + \cdots + (\underline{x}, \underline{a}_k)\underline{a}_k \ ,$$

and we may verify linearity as above.)

1.1.5 Example. Let A be a real $m \times n$-matrix. In this example we write elements from $\mathbb{R}^n$ and $\mathbb{R}^m$ as columns where necessary. Define a map $L_A : \mathbb{R}^n \to \mathbb{R}^m$ by

$$L_A(\underline{x}) := A\underline{x}$$

(on the right hand side the vector $\underline{x}$ is written as a column). This map is linear: let $\underline{x}$ and $\underline{y}$ be two elementens from $\mathbb{R}^n$ (written as columns), then it follows from the rules of matrix multiplication that for all scalars α and β it holds that:

$$A(\alpha\underline{x} + \beta\underline{y}) = \alpha A\underline{x} + \beta A\underline{y} \ ,$$

establishing the linearity of L_A. When for example

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

then the image of the vector $(3, 1, 1)$ from $\mathbb{R}^3$ equals $(4, 4)$ from $\mathbb{R}^2$; the image of $5 \cdot (3, 1, 1)$ therefore is $5 \cdot (4, 4) = (20, 20)$.

1.1.6 Example. For every vector space V the map $\mathcal{I} : V \to V$ given by $\mathcal{I}\underline{v} = \underline{v}$ is a linear map, the so called *identity map* or just *identity*. The trivial verification is left to the reader.

When both V and W are real or both complex vector spaces, then map $\mathcal{O} : V \to W$ given by $\mathcal{O}\underline{v} = \underline{0}$ is linear map, the so called *null map*.

1.1.7 Example. Differentiation is a linear map. Write $f' = Df$, then the following are elementary properties of the derivative

$$\begin{aligned} D(f+g) &= Df + Dg \ , \\ D(\alpha f) &= \alpha Df \ . \end{aligned}$$

There is a minor problem here: D is linear but between which vector spaces? Usually we make a choice depending on the circumstances, for instance D maps the space of polynomials to itself, or the space of differentiable functions to the space of all functions on $\mathbb{R}$.

1.1.8 In what follows we will show that compositions, inverses, sums and scalar multiples of linear maps are again linear maps. But first we have to define what exactly is meant by these maps and when they exist.

1.1.9 (Composition or product) Let $\mathcal{A} : V \to W$ and $\mathcal{B} : U \to V$ be linear maps. Take a vector $\underline{u} \in U$. Applying first $\mathcal{B}$, we get $\mathcal{B}\underline{u} \in V$. Now we may apply $\mathcal{A}$ and get $\mathcal{A}(\mathcal{B})\underline{u} \in W$, in a diagram:

$$\begin{array}{ccccc} U & \xrightarrow{\mathcal{B}} & V & \xrightarrow{\mathcal{A}} & W \\ \underline{u} & \longrightarrow & \mathcal{B}\underline{u} & \longrightarrow & \mathcal{A}(\mathcal{B}\underline{u}) \end{array} \tag{1.1}$$

1.1.10 Definition. Suppose $\mathcal{A} : V \to W$ and $\mathcal{B} : U \to V$ are linear maps. Then the *composition* $\mathcal{A} \circ \mathcal{B}$ or $\mathcal{A}\ \mathcal{B} : U \to W$ is defined by

$$(\mathcal{A}\mathcal{B})\underline{u} = \mathcal{A}(\mathcal{B}\underline{u}) \text{ for all } \underline{u} \in U \ .$$

1.1.11 Property. The composition is linear:

$$\begin{aligned} (\mathcal{A}\mathcal{B})(\underline{u}_1 + \underline{u}_2) &= \mathcal{A}(\mathcal{B}(\underline{u}_1 + \underline{u}_2)) = \mathcal{A}(\mathcal{B}\underline{u}_1 + \mathcal{B}\underline{u}_2) \\ &= (\mathcal{A}\mathcal{B})\underline{u}_1 + (\mathcal{A}\mathcal{B})\underline{u}_2 \ ; \\ (\mathcal{A}\mathcal{B})(\alpha\underline{u}) &= \mathcal{A}(\mathcal{B}(\alpha\underline{u})) = \mathcal{A}(\alpha\mathcal{B}\underline{u}) = \alpha(\mathcal{A}\mathcal{B})\underline{u} \ . \end{aligned}$$

1.1.12 The notation $\mathcal{A}\mathcal{B}$ for the composition suggests that we may view the composition as a "product" of the linear maps $\mathcal{A}$ and $\mathcal{B}$. This is only partly correct. When $\mathcal{A}\mathcal{B}$ exists, then $\mathcal{B}\mathcal{A}$ does not necessarily, for example when U, V and W are three different vector spaces. But even if $\mathcal{A} : V \to V$, $\mathcal{B} : V \to V$ are linear. Then both $\mathcal{A}\mathcal{B}$ and $\mathcal{B}\mathcal{A}$ exist, but need not to be equal. This does not stop us however from using the words product and composition to denote the same thing.

1.1.13 Example. Let V be the vector space of infinitely often differentiable functions and define $\mathcal{A} : V \to V$ by

$$\mathcal{A}f\ (x) = xf(x)\ , \quad \mathcal{B}f\ (x) = f'(x)\ .$$

Verify that these indeed are linear maps. Then

$$\begin{aligned} \mathcal{A}\mathcal{B}f\ (x) &= xf'(x)\ , \\ \mathcal{B}\mathcal{A}f\ (x) &= f(x) + xf'(x)\ , \end{aligned}$$

so $\mathcal{A}\mathcal{B} \neq \mathcal{B}\mathcal{A}$ (substitute for example the function $f : x \mapsto 1$).

1.1.14 (Inverse map) If the linear map $\mathcal{A} : V \to W$ is a bijection, then the inverse map $\mathcal{A}^{-1} : W \to V$ is also linear. Indeed, if $\underline{v}, \underline{w} \in W$ and α, β are scalars, then there are vectors $\underline{x}, \underline{y} \in V$ with $\mathcal{A}\underline{x} = \underline{v}$ and $\mathcal{A}\underline{y} = \underline{w}$ (because $\mathcal{A}$ is also surjective of course); from the linearity of $\mathcal{A}$ we get that $\mathcal{A}(\alpha\underline{x} + \beta\underline{y}) = \alpha\underline{v} + \beta\underline{w}$, which implies $\mathcal{A}^{-1}(\alpha\underline{v} + \beta\underline{w}) = \alpha\underline{x} + \beta\underline{y}$. On the other hand $\underline{x} = \mathcal{A}^{-1}\underline{v}$ and $\underline{y} = \mathcal{A}^{-1}\underline{w}$, so we conclude:

$$\mathcal{A}^{-1}(\alpha\underline{v} + \beta\underline{w}) = \alpha\mathcal{A}^{-1}\underline{v} + \beta\mathcal{A}^{-1}\underline{w} \ .$$

1.1.15 Definition. For $\mathcal{A} : V \to V$, we define $\mathcal{A}^2$ by $\mathcal{A}\mathcal{A}$ and more generally $\mathcal{A}^n = \mathcal{A}^{n-1}\mathcal{A}$ for $n = 2, 3, \ldots$. We take for $\mathcal{A}^0$ the identity map $\mathcal{I}$. In case $\mathcal{A}$ is invertible (so if the map has an inverse), then we denote for positive integers n by $\mathcal{A}^{-n}$ the composite map $\left(\mathcal{A}^{-1}\right)^n$.

1.1.16 Definition. (Sum, scalar multiple) Let $\mathcal{A} : V \to W$ and $\mathcal{B} : V \to W$ be two linear maps. Then the *sum* $\mathcal{A} + \mathcal{B} : V \to W$ is defined by

$$(\mathcal{A} + \mathcal{B})\underline{x} = \mathcal{A}\underline{x} + \mathcal{B}\underline{x} \ .$$

If α is a scalar, then the *scalar multiple* $\alpha\mathcal{A} : V \to W$ is defined by

$$(\alpha\mathcal{A})\underline{x} = \alpha(\mathcal{A}\underline{x}) \ .$$

1.1.17 Property. The sum $\mathcal{A} + \mathcal{B}$ and the scalar multiple $\alpha\mathcal{A}$ are linear maps:

$$\begin{aligned}
(\mathcal{A} + \mathcal{B})(\underline{x} + \underline{y}) &= \mathcal{A}(\underline{x} + \underline{y}) + \mathcal{B}(\underline{x} + \underline{y}) \\
&= \mathcal{A}\underline{x} + \mathcal{A}\underline{y} + \mathcal{B}\underline{x} + \mathcal{B}\underline{y} \\
&= (\mathcal{A} + \mathcal{B})\underline{x} + (\mathcal{A} + \mathcal{B})\underline{y} \ ; \\
(\mathcal{A} + \mathcal{B})(\alpha\underline{x}) &= \mathcal{A}(\alpha\underline{x}) + \mathcal{B}(\alpha\underline{x}) = \alpha\mathcal{A}\underline{x} + \alpha\mathcal{B}\underline{x} \\
&= \alpha(\mathcal{A}\underline{x} + \mathcal{B}\underline{x}) = \alpha(\mathcal{A} + \mathcal{B})\underline{x} \ .
\end{aligned}$$

Linearity of $\alpha\mathcal{A}$ follows in a similar way.

1.1.18 Example. Consider the linear differential equation

$$y'' - xy' + 2y = \sin x \ .$$

We may view the left hand side as a linear map applied to the vector y. Indeed, let D again stand for differentiation and define

$$\mathcal{A}_{f(x)}y = f(x)y \ .$$

For every f the map $\mathcal{A}_{f(x)}$ is linear, verify this. The left hand side of the differential equation can now be written as

$$y'' - xy' + 2y = (D^2 + \mathcal{A}_{-x}D + \mathcal{A}_2)y$$

and $D^2 + \mathcal{A}_{-x}D + \mathcal{A}_2$ is a linear map.

We did not specify here from which vector space to which vector space the linear maps go, this depends a little bit on the particular situation. Often the relevant vector spaces are the functions that are infinitely often differentiable.

1.1.19 Since linear maps are of course also ordinary maps, we can talk about the image of a vector, or of the image of a subset of the vector space (notation: $\mathcal{A}(D)$ if $\mathcal{A} : V \to W$ is the linear map and D the subset), and about the (complete) inverse image of a subset. For a subset D of W, we denote the inverse image of D under $\mathcal{A}$ by $\mathcal{A}^{-1}(D)$ of $\mathcal{A}^{\leftarrow}(D)$. The first notation is most common, but may cause confusion with the notation for the inverse in case the map $\mathcal{A}$ is not invertible, to avoid this we usually add something like 'the complete inverse image $\mathcal{A}^{-1}(D)$ of D'.

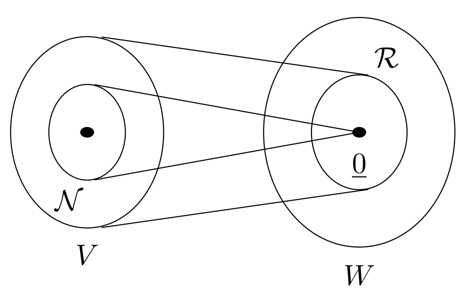
Associated to a linear map $\mathcal{A} : V \to W$ are two very important linear subspaces; the first is a generalization of the solution space of a homogeneous system of linear equations, the second is a generalization of the column space of a matrix.

1.1.20 Definition. Let $\mathcal{A} : V \to W$ be a linear map. Define

$$\mathcal{N} := \{\underline{v} \in V \mid \mathcal{A}\underline{v} = \underline{0}\} ,$$

$$\mathcal{R} := \{\mathcal{A}\underline{v} \in W \mid \underline{v} \in V\} .$$

$\mathcal{N}$ is called the *null space* of $\mathcal{A}$ and $\mathcal{R}$ the *range* (the null space is also called the *kernel*). The range can also be denoted by $\mathcal{A}(V)$.



1.1.21 $\mathcal{N}$ is a subset of V that always contains the zero vector $\underline{0}$; we verify that $\mathcal{N}$ is a linear subspace: if $\mathcal{A}\underline{v} = \underline{0}$ and $\mathcal{A}\underline{w} = \underline{0}$, and α and β are scalars, then

$$\mathcal{A}(\alpha\underline{v} + \beta\underline{w}) = \alpha\mathcal{A}\underline{v} + \beta\mathcal{A}\underline{w} = \underline{0} + \underline{0} = \underline{0} \ ,$$

so $\alpha\underline{v} + \beta\underline{w} \in \mathcal{N}$. Notice that the null space is precisely the inverse image $\mathcal{A}^{-1}\{\underline{0}\}$ of the origin under $\mathcal{A}$.

The range $\mathcal{R}$ is a subset of W and in fact a subspace of W; check this yourself.

1.1.22 Example. We determine the kernel and the range for example 1.1.5 with the given 2×3-matrix A. The null space of the linear map L_A consists of all vectors $\underline{x}$ satisfying

$$\begin{pmatrix} 1 & -1 & 2 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is a homogeneous system of linear equations with coefficient matrix A. The null space is the plane with equation $x_1 - x_2 + 2x_3 = 0$. The range of L_A consists of all vectors of the form

$$\begin{pmatrix} 1 & -1 & 2 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

that is, vectors of the form

$$x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

But this is exactly the span of the columns of A, that is, the column space. We conclude that the range equals $< (1, 1) >$.

We also determine the inverse image $L_A^{-1}(\ell)$ of the line ℓ with vector presentation

$$\ell : \underline{x} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

So we are looking for vectors $\underline{x}$ satisfying

$$A\underline{x} = \begin{pmatrix} 3 + 2\lambda \\ 2 + \lambda \end{pmatrix},$$

for certain λ. This means that we have to solve the system with augmented matrix

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 3+2\lambda \\ 1 & -1 & 2 & 2+\lambda \end{array}\right).$$

Using row reduction (Gaussian elimination) one sees that this system only has solutions for $\lambda = -1$. The solutions constitute the plane with equation $x_1 - x_2 + 2x_3 = 1$. Can you see from the mutual position of ℓ and the range $\mathcal{R}$ that computing the inverse image can be restricted to that of the vector $(1,1)$?

1.1.23 Example. The null space of the linear map L_A from example 1.1.5 consists of the vectors $\underline{x}$ that satisfy $L_A(\underline{x}) = \underline{0}$, so all solutions of the homogeneous system $A\underline{x} = \underline{0}$.

The range of L_A consists of all vectors of the form $L_A(\underline{x})$. If the columns of A are $\underline{a}_1, \ldots, \underline{a}_n$, then this is the set

$$\{x_1\underline{a}_1 + \cdots + x_n\underline{a}_n \mid x_1, \ldots, x_n \text{ arbitrary}\}\ .$$

This is the column space of A.

So null space and range generalise two notions from the matrix world.

1.1.24 Example. In $\mathbb{R}^2$ we consider the orthogonal projection $\mathcal{P}$ on a line (through the origin) $\ell =< \underline{a} >$, see example 1.1.4. Geometrically we see that the range, the collection of vectors that occur as an image, is ℓ itself. The null space consists of all vectors that are mapped to $\underline{0}$ and that is $\ell^{\perp}$. We can also see this from the formula $\mathcal{P}\underline{x} = (\underline{x}, \underline{a})\underline{a}$ (where $\underline{a}$ has length 1). The null space for example consists of all vectors $\underline{x}$ satisfying $(\underline{x}, \underline{a})\underline{a} = 0$, and since $\underline{a} \neq \underline{0}$ this means $(\underline{x}, \underline{a}) = 0$, which is the orthoplement (orthogonal complement) of $\underline{a}$.

1.1.25 We now connect the notions injectivity, surjectivity and inverse image to the notions null space and range of a linear map.

1.1.26 Theorem. *Consider a linear map $\mathcal{A} : V \to W$.*

1. $\mathcal{N} = \{\,\underline{0}\,\} \Leftrightarrow \mathcal{A}$ *is injective.*
2. $\mathcal{R} = W \Leftrightarrow \mathcal{A}$ *is surjective.*
3. *Let $\underline{b} \in \mathcal{R}$. Then there is a vector $\underline{p}$ satisfying $\mathcal{A}\underline{p} = \underline{b}$; we say that $\underline{p}$ is a* particular solution *of the vector equation $\mathcal{A}\underline{x} = \underline{b}$. Together all solutions of the vector equation $\mathcal{A}\underline{x} = \underline{b}$ are given by the set*

$$\underline{p} + \mathcal{N} = \{\underline{p} + \underline{n} \mid \underline{n} \in \mathcal{N}\}\ .$$

In particular, the equation $\mathcal{A}\underline{x} = \underline{b}$ has exactly one solution if $\mathcal{N} = \{\underline{0}\}$.

Proof.

1. $\Rightarrow$) We have to show that $\mathcal{A}\underline{x} = \mathcal{A}\underline{y}$ implies $\underline{x} = \underline{y}$. If $\mathcal{A}\underline{x} = \mathcal{A}\underline{y}$, then $\mathcal{A}(\underline{x} - \underline{y}) = \mathcal{A}\underline{x} - \mathcal{A}\underline{y} = \underline{0}$, so that $\underline{x} - \underline{y} \in \mathcal{N}$. Since $\mathcal{N}$ contains only the zero vector, we conclude that $\underline{x} - \underline{y} = \underline{0}$ or $\underline{x} = \underline{y}$.
   $\Leftarrow$) If $\underline{x} \in \mathcal{N}$, then $\mathcal{A}\underline{x} = \mathcal{A}\underline{0}$. Injectivity of $\mathcal{A}$ now implies $\underline{x} = \underline{0}$, so the only element of $\mathcal{N}$ is the zero vector.

2. This is the definition of surjectivity.

3. The first part of the statement is trivial. For the second part: Let $\underline{p}$ be a particular solution. For every $\underline{n} \in \mathcal{N}$ we have $\mathcal{A}(\underline{p}+\underline{n}) = \mathcal{A}\underline{p}+\mathcal{A}\underline{n} = \underline{b} + \underline{0} = \underline{b}$; so $\underline{p} + \underline{n}$ is also a solution. Conversely, if $\underline{q}$ is a solution, then $\mathcal{A}(\underline{q} - \underline{p}) = \mathcal{A}\underline{q} - \mathcal{A}\underline{p} = \underline{b} - \underline{b} = \underline{0}$, so that $\underline{q} - \underline{p} \in \mathcal{N}$. Since $\underline{q} = \underline{p} + (\underline{q} - \underline{p})$, $\underline{q}$ is the sum of $\underline{p}$ and a vector from the null space. All solutions therefore form the set $\{\underline{p} + \underline{n} \mid \underline{n} \in \mathcal{N}\}$.

$\square$

1.1.27 Short hand notation for the set $\{\underline{p} + \underline{n} \mid \underline{n} \in \mathcal{N}\}$ is $\underline{p} + \mathcal{N}$. Such a set is also called a *coset*.

1.1.28 Let $\mathcal{A} : V \to W$ be a linear map and consider the equation $\mathcal{A}\underline{x} = \underline{b}$. The corresponding homogeneous equation is (by definition) the equation $\mathcal{A}\underline{x} = \underline{0}$. The solutions of this equation form the null space of $\mathcal{A}$. So theorem 1.1.26 says that we get all solutions of the vector equation $\mathcal{A}\underline{x} = \underline{b}$ by taking *one* particular solution and adding to that all solutions of the corresponding homogeneous equation.

This property is frequently used in the solution of linear differential equations, a differential equation of the form $\mathcal{A}y = f$.

1.1.29 (Specifying a linear map) When we know the images $\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_n$ then the image of any linear combination $x_1\underline{a}_1 + \cdots + x_n\underline{a}_n$ follows because of linearity:

$$\mathcal{A}(x_1\underline{a}_1 + \cdots + x_n\underline{a}_n) = x_1\mathcal{A}\underline{a}_1 + \cdots + x_n\mathcal{A}\underline{a}_n.$$

In particular it is possible to determine the image of any vector when the images of the vectors in a basis are known. This simple observation leads to the following theorem.

1.1.30 Theorem. *Let V and W be vector spaces, $\{\underline{a}_1, \ldots, \underline{a}_n\}$ a basis for V and $\underline{w}_1, \ldots, \underline{w}_n$ an n–tuple of vectors in W. Then there exists a unique linear map $\mathcal{A} : V \to W$ satisfying $\mathcal{A}\underline{a}_i = \underline{w}_i$ for $i = 1, \ldots, n$.*

Proof. Take an arbitrary vector $\underline{x} \in V$. This vector can be written uniquely as

$$\underline{x} = \sum_{i=1}^{n} x_i \underline{a}_i \ .$$

This means that for the image vector $\mathcal{A}\underline{x}$ we must have

$$\mathcal{A}\underline{x} = \sum_{i=1}^{n} x_i \mathcal{A}\underline{a}_i = \sum_{i=1}^{n} x_i \underline{w}_i \ . \qquad *$$

It follows that this vector is uniquely determined by the vector $\underline{x}$. We may simply verify that when we use the relation $(*)$ as the definition of a map $\mathcal{A} : V \to W$, then this map $\mathcal{A}$ is linear. □

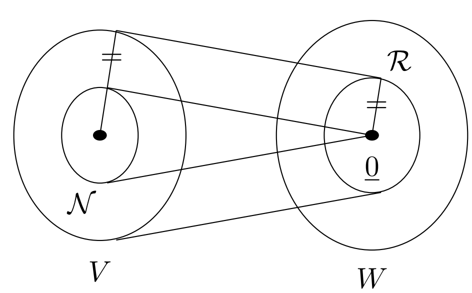
1.1.31 Theorem. *Consider a linear map $\mathcal{A} : V \to W$ with $V =< \underline{a}_1, \ldots, \underline{a}_n >$ (see B.2.3). Then $\mathcal{R} =< \mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_n >$.*

Proof. $\mathcal{R}$ consists of all vectors $\mathcal{A}\underline{x}$ with $\underline{x} \in V$. Such an $\underline{x}$ is a linear combination of $\underline{a}_1, \ldots, \underline{a}_n$ hence such an $\mathcal{A}\underline{x}$ is a linear combination of $\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_n$. □

1.1.32 A similar characterisation of $\mathcal{N}$ is not as easy. We do however have the following important result:

1.1.33 Theorem. (Dimension Theorem) *Let $\mathcal{A} : V \to W$ be a linear map with $\dim(V) < \infty$. Then*

$$\dim(V) = \dim(\mathcal{N}) + \dim(\mathcal{R}) \ .$$



Proof. Suppose $\dim(V) = n$. Since $\mathcal{N} \subset V$, we have $\dim \mathcal{N} = p \leq n$. Choose a basis $\{\underline{a}_1, \ldots, \underline{a}_p\}$ for $\mathcal{N}$ and complete this with $\underline{b}_{p+1}, \ldots, \underline{b}_n$ to a *basis* for V:

$$V =< \underline{a}_1, \ldots, \underline{a}_p,\ \underline{b}_{p+1}, \ldots, \underline{b}_n > \ .$$

Then

$$\begin{aligned}\mathcal{R} &=< \mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_p, \mathcal{A}\underline{b}_{p+1}, \ldots, \mathcal{A}\underline{b}_n > \\ &=< \mathcal{A}\underline{b}_{p+1}, \ldots, \mathcal{A}\underline{b}_n >\end{aligned}$$

since $\mathcal{A}\underline{a}_1 = \cdots = \mathcal{A}\underline{a}_p = \underline{0}$. We are finished when we can show that the vectors $\mathcal{A}\underline{b}_{p+1}, \ldots, \mathcal{A}\underline{b}_n$ are independent, because now $\dim(\mathcal{R}) = n - p$ and hence $n = \dim(V) = p + (n - p) = \dim \mathcal{N} + \dim \mathcal{R}$.
To prove independence we assume

$$\alpha_{p+1}\mathcal{A}\underline{b}_{p+1} + \cdots + \alpha_n \mathcal{A}\underline{b}_n = \underline{0} \ ,$$

so that succesively

$$\mathcal{A}(\alpha_{p+1}\underline{b}_{p+1} + \cdots + \alpha_n \underline{b}_n) = \underline{0} \ ,$$

$$\alpha_{p+1}\underline{b}_{p+1} + \cdots + \alpha_n \underline{b}_n \in \mathcal{N} \ ,$$

$$\alpha_{p+1}\underline{b}_{p+1} + \cdots + \alpha_n \underline{b}_n = \alpha_1 \underline{a}_1 + \cdots + \alpha_p \underline{a}_p \text{ for certain scalars } \alpha_1, \ldots, \alpha_p \ ,$$

$$-\alpha_1 \underline{a}_1 - \cdots - \alpha_p \underline{a}_p + \alpha_{p+1}\underline{b}_{p+1} + \cdots + \alpha_n \underline{b}_n = \underline{0} \ .$$

Since $\{\underline{a}_1, \ldots, \underline{a}_p, \underline{b}_{p+1}, \ldots, \underline{b}_n\}$ is a basis it follows from this that $\alpha_1 = \cdots = \alpha_p = \alpha_{p+1} = \cdots = \alpha_n = 0$; the system $\mathcal{A}\underline{b}_{p+1}, \ldots, \mathcal{A}\underline{b}_n$ is therefore independent. □

1.1.34 Example. The range of the orthogonal projection $\mathcal{P}$ on a subspace W of the inner product space V (see example 1.1.4) obviously equals W, while the null space consists of all vectors perpendicular (orthogonal) to W, so $\mathcal{N} = W^{\perp}$. The dimension theorem above therefore implies that $\dim V = \dim W + \dim W^{\perp}$.

1.1.35 (Inverse) To conclude we investigate when a linear map $\mathcal{A} : V \to W$ with $\dim(V) < \infty$ is invertible. The result is as follows:

1.1.36 Theorem. *A linear map* $\mathcal{A} : V \to W$ *with* $\dim(V) < \infty$ *has an inverse if and only if* $\dim(V) = \dim(W)$ *and* $\mathcal{N} = \{\underline{0}\}$.

Proof. When $\mathcal{A}$ has an inverse, then the map is a bijection and hence injective and surjective. From Theorem 1.1.26 it follows that $\mathcal{N} = \{\underline{0}\}$ and $W = \mathcal{R}$. The dimension theorem, $\dim(V) = \dim(\mathcal{N}) + \dim(\mathcal{R})$, then implies $\dim(V) = \dim(\mathcal{R}) = \dim(W)$. Invertibility therefore implies $\mathcal{N} = \{\underline{0}\}$ and $\dim V = \dim W$.

Conversely, if $\mathcal{N} = \{\underline{0}\}$ and $\dim V = \dim W$, then $\mathcal{A}$ is invertible, for if $\mathcal{N} = \{\underline{0}\}$ then from the dimension theorem it follows that $\dim V = \dim(\mathcal{R})$; together with the fact that $\dim V = \dim W$ this implies $\dim(\mathcal{R}) = \dim W$. Apparently $\mathcal{R} = W$. From $\mathcal{N} = \{\underline{0}\}$, $\mathcal{R} = W$ and Theorem 1.1.26 we finally get that $\mathcal{A}$ is injective and surjective, and hence invertible. □

1.1.37 Check for yourself that in this theorem we may replace the condition $\mathcal{N} = \{\underline{0}\}$ by the condition $\mathcal{R} = W$.

1.2 Matrices of linear maps I

1.2.1 In this section we study linear maps $\mathbb{K}^n \to \mathbb{K}^m$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ or even $\mathbb{Q}$ or a more exotic field. We will use matrices to describe such maps.

1.2.2 (The matrix of a linear map) If necessary, look up in linear algebra 1 the definition of matrix multiplication. Consider an $m \times n$ matrix A:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

For every n-tuple $\underline{x} = (x_1, x_2, \dots, x_n)$ we define $\mathcal{A}\underline{x} = (y_1, y_2, \dots, y_m)$ by

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} .$$

In example 1.1.5 we saw that this rule defines a linear map $\mathcal{A} : \mathbb{K}^n \to \mathbb{K}^m$.

Let $\underline{e}_1, \dots, \underline{e}_n$ be the standard basis of $\mathbb{K}^n$ and let $\underline{k}_1, \dots, \underline{k}_n$ be the columns of the matrix A. We see by substitution that $\mathcal{A}\underline{e}_1 = A\underline{e}_1 = \underline{k}_1, \mathcal{A}\underline{e}_2 = \underline{k}_2, \dots, \mathcal{A}\underline{e}_n = \underline{k}_n$: *the columns of the matrix A are the images under $\mathcal{A}$ of the basis vectors $\underline{e}_1, \dots, \underline{e}_n$.*

Now start with an arbitrary linear map $\mathcal{A} : \mathbb{K}^n \to \mathbb{K}^m$ and write a vector $\underline{x} \in \mathbb{K}^n$ as a linear combination of the standard basis vectors: $\underline{x} = x_1\underline{e}_1 + \cdots + x_n\underline{e}_n$. Then $\mathcal{A}\underline{x} = x_1\mathcal{A}\underline{e}_1 + \cdots + x_n\mathcal{A}\underline{e}_n$ because of the linearity of $\mathcal{A}$. Collect the n vectors $\mathcal{A}\underline{e}_1, \ldots, \mathcal{A}\underline{e}_n$ as columns in an $m \times n$–matrix A. Now $x_1\mathcal{A}\underline{e}_1 + \cdots + x_n\mathcal{A}\underline{e}_n$ equals the matrix product $A\underline{x}$. So we found a new (and very important) use for matrices:

1.2.3 Theorem. *Every linear map $\mathcal{A} : \mathbb{K}^n \to \mathbb{K}^m$ is determined by an $m \times n$-matrix A whose columns are $\mathcal{A}\underline{e}_1, \ldots, \mathcal{A}\underline{e}_n$. The image of the vector $\underline{x}$ under $\mathcal{A}$ can be computed as the matrix product $A\underline{x}$.*

1.2.4 Definition. The matrix A is called *the matrix of the linear map* $\mathcal{A}$.

1.2.5 (Computing the matrix) As a consequence of Theorem 1.1.30 a linear map $\mathcal{A} : \mathbb{K}^n \to \mathbb{K}^m$ is uniquely determined by the images of a basis $\underline{a}_1, \ldots, \underline{a}_n$ for $\mathbb{K}^n$. If $\underline{e}_1, \ldots, \underline{e}_n$ is this basis, then we may just write down the matrix using the previous theorem, but if we start with another basis, we need to do some computations. Let $\mathcal{A}$ be a linear map and consider two ordered pairs of row-vectors, on the left a vector in $\mathbb{K}^n$, and on the right its image in $\mathbb{K}^m$:

$$(\underline{a}, \mathcal{A}\underline{a}) \ ,$$
$$(\underline{b}, \mathcal{A}\underline{b}) \ .$$

Adding these two row-vectors we get

$$(\underline{a} + \underline{b}, \mathcal{A}\underline{a} + \mathcal{A}\underline{b}) \ .$$

Since $\mathcal{A}$ is linear, the right hand side equals $\mathcal{A}(\underline{a} + \underline{b})$; so addition gives us again a row with a vector on the left and its image on the right. The same applies to scalar multiplication: multiplying a row with α we find

$$(\alpha\underline{a}, \ \alpha\mathcal{A}\underline{a})$$

and because of linearity $\alpha\mathcal{A}\underline{a} = \mathcal{A}(\alpha\underline{a})$, so this again gives a row with on the left a vector and on the right its image. So if we perform elementary row operations (adding rows and multiplying them by non-zero scalars) in such a system we keep the property that on the right we have the image of the vector on the left.

1.2.6 Example. The linear map $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$\mathcal{A}(-1, 0, 1) = (-4, 2, 4), \ \mathcal{A}(1, 1, 0) = (1, -1, -1), \ \mathcal{A}(0, 1, 2) = (-5, 4, 4) \ .$$

We put these data in three (vector, image)-rows, for clarity we put a vertical bar in between the vector and its image:

$$\left(\begin{array}{rrr|rrr} -1 & 0 & 1 & -4 & 2 & 4 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 0 & 1 & 2 & -5 & 4 & 4 \end{array}\right) .$$

Now we perform row-reduction, and obtain the (row-reduced) normal form

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 2 & 1 & -3 \\ 0 & 1 & 0 & -1 & -2 & 2 \\ 0 & 0 & 1 & -2 & 3 & 1 \end{array}\right) .$$

Now we have in our rows on the left the standard basis vectors $\underline{e}_1, \underline{e}_2, \underline{e}_3$ and on the right their images $\mathcal{A}\underline{e}_1, \mathcal{A}\underline{e}_2, \mathcal{A}\underline{e}_3$. According to Theorem 1.2.3 the matrix of $\mathcal{A}$ is

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

It is easy to check whether we made an error by using this matrix to compute the images of the vectors $(-1, 0, 1)$, $(1, 1, 0)$ and $(0, 1, 2)$.

1.2.7 (Connection with systems of linear equations) Consider the following system of linear equations

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

with $m \times n$–coefficient matrix A. This matrix determines a linear map $\mathcal{A} : \mathbb{K}^n \to \mathbb{K}^m$ and we see that the system of equations can be written as a vector equation

$$\mathcal{A}\underline{x} = \underline{b} .$$

The range $\mathcal{R}$ of $\mathcal{A}$ is the space $< \mathcal{A}\underline{e}_1, \ldots, \mathcal{A}\underline{e}_n >$ according to Theorem 1.1.31; this is therefore the column space of the matrix A. The equation $\mathcal{A}\underline{x} = \underline{b}$ has a solution (at least one) if and only if $\underline{b} \in \mathcal{R}$, so if and only if $\underline{b}$ is contained in the column space of the matrix. Because of Theorem 1.1.26 we find all solutions of the equation $\mathcal{A}\underline{x} = \underline{b}$ as the sum of one particular solution and all vectors from the null space.

The null space consists of all vectors $\underline{x}$ with $\mathcal{A}\underline{x} = \underline{0}$, so all solutions of the homogeneous system $A\underline{x} = \underline{0}$.

1.2.8 In the rest of this section we discuss the matrix of the composition and the sum of two linear maps, and of the inverse, whenever it makes sense.

1.2.9 (Matrix composition) Let $\mathcal{A} : \mathbb{K}^m \to \mathbb{K}^q$ be a linear map with $q \times m$–matrix A and let $\mathcal{B} : \mathbb{K}^p \to \mathbb{K}^m$ a map with $m \times p$–matrix B. From the way the matrix of $\mathcal{A}$ is constructed we have that for every vector $\underline{x} \in \mathbb{K}^m$ that the image $\mathcal{A}\underline{x}$ is equal to the matrix product $A\underline{x}$. Same for $\mathcal{B}$.

To the composite map $\mathcal{AB} : \mathbb{K}^p \to \mathbb{K}^q$ now belongs a $q \times p$–matrix, whose j–th column is equal to $(\mathcal{AB})\underline{e}_j$. We now have

$$(\mathcal{AB})\underline{e}_j = \mathcal{A}(\mathcal{B}\underline{e}_j) = \mathcal{A}(B\underline{e}_j) = A(B\underline{e}_j) = (AB)\underline{e}_j,$$

but this is exactly the j–the column of AB. We conclude that the matrix of the product $\mathcal{AB}$ is the matrix product AB. So we see

1.2.10 Theorem. (Matrix composition) *Let A and B be two matrices such that AB exists and let $\mathcal{A}$ and $\mathcal{B}$ be the corresponding linear maps. Then AB is the matrix of the product map $\mathcal{AB}$.*

The following properties are proved in a similar way:

1.2.11 Theorem. (Matrix of sum and scalar multiple) *Let A and B be two matrices of the same size and let $\mathcal{A}$ and $\mathcal{B}$ be the corresponding linear maps. Then $A + B$ is the matrix of sum $\mathcal{A} + \mathcal{B}$ and for every scalar α, the matrix of the linear map $\alpha\mathcal{A}$ is αA.*

1.2.12 (Matrix of the inverse) At the end of this section we consider the connection between the inverse of a linear map and the inverse of a matrix. If $\mathcal{A} : \mathbb{K}^n \to \mathbb{K}^m$ has an inverse, then $m = n$ because of Theorem 1.1.36. The matrix A of $\mathcal{A}$ is therefore an $n \times n$–matrix. The inverse map, $\mathcal{B}$ say, has matrix B. Since $\mathcal{AB} = \mathcal{BA} = \mathcal{I}$ it follows from Theorem 1.2.10 that $AB = BA = I$. Matrix A is therefore invertible, moreover $B = A^{-1}$.

Conversely, when A is an invertible $n \times n$–matrix, then the corresponding linear map $\mathcal{A}$ has an inverse: Indeed, A has rank n so the system $A\underline{x} = \underline{0}$ has only the zero vector as a solution: $\mathcal{N} = \{\underline{0}\}$. From Theorem 1.1.36 it now follows that $\mathcal{A}$ is invertible. We have found:

1.2.13 Theorem. (Matrix of the inverse) *Let $\mathcal{A} : \mathbb{K}^n \to \mathbb{K}^n$ be a linear map with matrix A. Then the map $\mathcal{A}$ has an inverse if and only if the matrix A has an inverse. If $\mathcal{A}$ has an inverse, then the matrix of $\mathcal{A}^{-1}$ is equal to A^{-1}.*

1.2.14 When we introduced the inverse of a matrix it was mentioned without proof that $AB = I$ implies $BA = I$ for square matrices A and B. In the context of linear maps we can easily prove this. Let A and B be $n \times n$ matrices such that $AB = I$, and let $\mathcal{A}$ and $\mathcal{B}$ be the corresponding linear maps of $\mathbb{K}^n$ to $\mathbb{K}^n$. Then it follows from Theorem 1.2.10 that $\mathcal{A}\mathcal{B} = \mathcal{I}$, in other words $\mathcal{A}\mathcal{B}\underline{x} = \underline{x}$ for all $\underline{x} \in \mathbb{K}^n$. Now if $\underline{x}$ is in the null space of $\mathcal{B}$ then it follows immediately that $\underline{x} = \underline{0}$. The null space of $\mathcal{B}$ therefore consists solely of the zero vector and with Theorem 1.1.36 we conclude that $\mathcal{B}$ has an inverse. It follows that $\mathcal{A}$ is this inverse and therefore also $\mathcal{B}\mathcal{A} = \mathcal{I}$. Going back to matrices we get $BA = I$. The matrices A and B are therefore each others inverse. In short: $AB = I \Leftrightarrow BA = I$.

1.2.15 The considerations above give us an alternative way to compute the inverse of a square (invertible) matrix: the matrix comes with a linear map, compute the matrix of the inverse map. We will show how to do this by giving two examples.

1.2.16 **Example.** Consider the matrix

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 4 & -3 & 8 \end{pmatrix} .$$

For the corresponding linear map $\mathcal{A}$ we have $\mathcal{A}\underline{e}_1 = (1, 0, 4)$, $\mathcal{A}\underline{e}_2 = (0, 1, -3)$ and $\mathcal{A}\underline{e}_3 = (3, 2, 8)$. For the inverse map $\mathcal{A}^{-1}$ it follows that $\mathcal{A}^{-1}(1, 0, 4) = \underline{e}_1$, $\mathcal{A}^{-1}(0, 1, 3) = \underline{e}_2$ and $\mathcal{A}^{-1}(3, 2, 8) = \underline{e}_3$. We now determine the matrix of $\mathcal{A}^{-1}$ exactly as in 1.2.5:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & -3 & 0 & 1 & 0 \\ 3 & 2 & 8 & 0 & 0 & 1 \end{array}\right).$$

Row reduction yields the normal form

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 7 & 4 & -2 \\ 0 & 1 & 0 & -\frac{9}{2} & -2 & \frac{3}{2} \\ 0 & 0 & 1 & -\frac{3}{2} & -1 & \frac{1}{2} \end{array}\right)$$

from which we see the inverse

$$A^{-1} = \frac{1}{2} \begin{pmatrix} 14 & -9 & -3 \\ 8 & -4 & -2 \\ -4 & 3 & 1 \end{pmatrix} .$$

1.2.17 Example. Consider the matrix

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 4 & -3 & 6 \end{pmatrix} .$$

For the inverse $\mathcal{A}^{-1}$ of the corresponding linear map $\mathcal{A}$, (provided it exists), we must have $\mathcal{A}^{-1}(1,0,4) = \underline{e}_1$, $\mathcal{A}^{-1}(0,1,-3) = \underline{e}_2$, $\mathcal{A}^{-1}(3,2,6) = \underline{e}_3$. So

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & -3 & 0 & 1 & 0 \\ 3 & 2 & 6 & 0 & 0 & 1 \end{array}\right) .$$

Partial row reduction gives

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 & -2 & 1 \end{array}\right) .$$

Apparently the columns of A are dependent. So there is no inverse matrix.

1.3 Matrices of linear maps II

1.3.1 As before $\mathbb{K}$ will stand for $\mathbb{R}$ or $\mathbb{C}$. In the previous section we saw that linear maps from $\mathbb{K}^n$ to $\mathbb{K}^m$ can be described using matrices. Such a description is also possible for other vector spaces, but we then need to choose a basis, so that we can work with coordinates. In this section we discuss this.

- we analyse the effect on the coordinates of a vector when changing to another basis,
- we describe linear maps using matrices depending on the choice of a basis and
- we deduce how these matrices transform under a change of basis.

These techniques will then be used (in subsequent sections) to search for bases with respect to which the matrix of a linear map $V \to W$ is "simple".

1.3.2 (Coordinates) In an n-dimensional vector space V we choose a basis $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$. Every vector $\underline{x} \in V$ can now be expressed in exactly one way as a linear combination of the basis vectors

$$\underline{x} = x_1\underline{a}_1 + x_2\underline{a}_2 + \cdots + x_n\underline{a}_n \; .$$

The numbers $x_1, \ldots, x_n$ are the coordinates of $\underline{x}$ with respect to the basis α and $(x_1, \ldots, x_n)$ is the coordinate vector of $\underline{x}$ with respect to the basis α. The map associating to each vector $\underline{x}$ the corresponding coordinate vector is therefore a bijection. It is clear that the coordinates depend on the choice of the basis α.

With respect to a fixed basis α, the coordinates of the sum $\underline{x}+\underline{y}$ are the sum of the coordinates of $\underline{x}$ and of $\underline{y}$ and the coordinates of $\alpha\underline{x}$ are precisely α times the coordinates of $\underline{x}$. But this means:

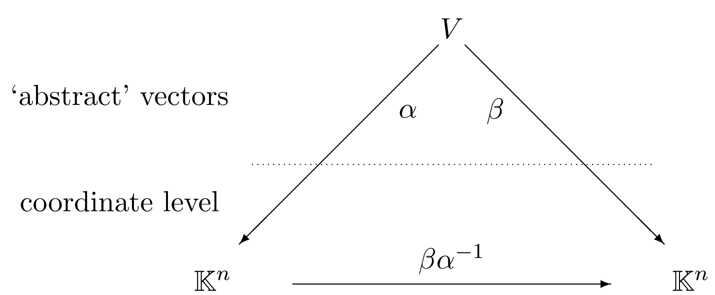
1.3.3 Property. If we choose a basis α in an n-dimensional vector space V, then the map sending each vector $\underline{x}$ to its coordinates with respect to this basis is an invertible linear map from V to $\mathbb{K}^n$.

We will usually denote this map also by α. This should not lead to confusion: $\alpha(\underline{x})$ is the coordinaat vector of the vector $\underline{x} \in V$ with respect to the basis α.

1.3.4 (Change of basis) We now consider an n-dimensional vector space V and choose for V two bases:

$$\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\} \quad \text{and} \quad \beta = \{\underline{b}_1, \ldots, \underline{b}_n\}.$$

So now with every $\underline{x} \in V$ there correspond two sets of coordinates: $\alpha(\underline{x})$ with respect to the basis α and $\beta(\underline{x})$ with respect to the basis β.



It should now be clear how the α-coordinates and the β-coordinates of $\underline{x}$ are related: we start with the α-coordinates, viewed as a vector in $\mathbb{K}^n$, apply the map α^{-1} giving us $\underline{x} \in V$, then we apply the map β giving us $\beta(\underline{x}) \in \mathbb{K}^n$.

1.3.5 Definition. (Coordinate transformation (map)) Let α and β be bases of an n-dimensional vector space V. The linear map $\beta\,\alpha^{-1} : \mathbb{K}^n \to \mathbb{K}^n$ is called the *coordinate transformation (map)* from α to β.

1.3.6 Coordinate transformation is a linear map from $\mathbb{K}^n$ to itself, hence corresponds to an $n \times n$–matrix that we denote by ${}_\beta S_\alpha$ whose columns are the images of $\underline{e}_1, \ldots, \underline{e}_n$. Now $(\beta\alpha^{-1})(\underline{e}_i) = \beta(\alpha^{-1}(\underline{e}_i)) = \beta(\underline{a}_i)$. So the columns of this matrix are the β–coordinate vectors of the α–basis vectors. The map $\beta\,\alpha^{-1}$ transforms by construction α–coordinates to β–coordinates. This transformation can therefore be computed using multiplication with the matrix ${}_\beta S_\alpha$. Summarizing:

1.3.7 Theorem. *Let α and β be bases of an n-dimensional vector space V and let ${}_\beta S_\alpha$ be the matrix of $\beta\,\alpha^{-1}$. If $\underline{x}$ is the α–coordinate vector of a vector $\underline{v} \in V$, then the β–coordinate vector of $\underline{v}$ is equal to the product ${}_\beta S_\alpha\,\underline{x}$*

1.3.8 Definition. (Transition matrix) The matrix ${}_\beta S_\alpha$ is called the *transition matrix* of the basis α to the basis β.

1.3.9 It is of course also possible to change β–coordinates to α–coordinates. This is done with the matrix ${}_\alpha S_\beta$. The product matrix ${}_\alpha S_\beta\ {}_\beta S_\alpha$ transforms α–coordinates of a vector $\underline{x}$ to β–coordinates and then back to the α–coordinates of $\underline{x}$. So

$$ {}_\alpha S_\beta\ {}_\beta S_\alpha = I\,, \quad {}_\alpha S_\beta = {}_\beta S_\alpha^{-1}. $$

If a third basis γ enters the game, then we may transform α–coordinates to β–coordinates and these subsequently to γ–coordinates. The result of this complicated game is of course that we just changed α–coordinates to γ–coordinates, in a formula:

$$ {}_\gamma S_\beta\,{}_\beta S_\alpha = {}_\gamma S_\alpha $$

(this identity also follows by applying Theorem 1.2.10 to $(\gamma\beta^{-1})(\beta\alpha^{-1}) = \gamma\alpha^{-1}$).

1.3.10 Example. We consider the vector space V of real polynomials of degree at most 2. We consider the bases $\alpha:\ \{1, x, x^2\}$ and $\beta:\ \{x-1, x^2-1, x^2+1\}$. We can easily express the basis vectors of β in the basis vectors of α:

$$ \begin{aligned} x-1 &= (-1)\cdot 1 + 1\cdot x + 0\cdot x^2 \\ x^2-1 &= (-1)\cdot 1 + 0\cdot x + 1\cdot x^2 \\ x^2+1 &= \quad 1\cdot 1 + 0\cdot x + 1\cdot x^2 \quad . \end{aligned} $$

We know therefore the α–coordinates of the vectors of β. So

$$ {}_\alpha S_\beta = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}. $$

The transition matrix ${}_\beta S_\alpha$ is the inverse of this matrix. We find

$$ {}_\beta S_\alpha = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix} . $$

The first column of this matrix should consist of the β–coordinates of the first basisvector of α. These β–coordinates are $(0, -\frac{1}{2}, \frac{1}{2})$ corresponding to the vector $0\,(x-1) - \frac{1}{2}(x^2-1) + \frac{1}{2}(x^2+1) = 1$, so that is correct. Verify that the second column consists of the β–coordinates of x and the third column of the β–coordinates of x^2.

How do we determine the β–coordinates of the vector $2x^2 - 3x + 4$? The α-coordinates of this vector are $(4, -3, 2)$. We transform them to β–coordinates using the transition matrix ${}_\beta S_\alpha$:

$$ {}_\beta S_\alpha \begin{pmatrix} 4 \\ -3 \\ 2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ \frac{1}{2} \\ \frac{3}{2} \end{pmatrix} . $$

We verify the result:

$$ -3(x-1) + \frac{1}{2}(x^2-1) + \frac{3}{2}(x^2+1) = 2x^2 - 3x + 4 . $$

Correct.

1.3.11 It is important to distinguish between calculating with vectors, so with elements of the vector space V, and calculating with coordinates, so with sequences of numbers from $\mathbb{K}^n$. In the example that was clear: the vectors are polynomials, the coordinate vectors are sequences of real or complex numbers.

This distinction becomes more problematic if our vector space is $\mathbb{K}^n$ itself. The sequence that represents a vector from $\mathbb{K}^n$ can also be seen as sequence of coordinates, w.r.t. the standard basis $\varepsilon : \{\underline{e}_1, \ldots, \underline{e}_n\}$:

$$ (1,2,3) = 1\underline{e}_1 + 2\underline{e}_2 + 3\underline{e}_3 . $$

1.3.12 **Example.** We consider in $\mathbb{R}^3$ two bases:

$$ \begin{aligned} \alpha &= \{(1,0,2),\ (-1,1,0),\ (0,-2,1)\} , \\ \beta &= \{(0,1,1),\ (1,2,-1),\ (1,0,1)\} . \end{aligned} $$

We are looking for the transition matrix ${}_\beta S_\alpha$. So with this matrix we transform α–coordinates into β–coordinates. The columns of this matrix are the β–coordinates of the vectors of α.

One way to proceed is to solve this problem directly. We write each of the vectors of α as a linear combination of the vectors of β. This means we have to solve three systems of equations in three unknowns, that we may combine into:

$$\left(\begin{array}{rrr|rrr} 0 & 1 & 1 & 1 & -1 & 0 \\ 1 & 2 & 0 & 0 & 1 & -2 \\ 1 & -1 & 1 & 2 & 0 & 1 \end{array}\right) .$$

Row reduction gives

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{4} & 0 & -\frac{3}{4} \\ 0 & 0 & 1 & \frac{5}{4} & -1 & \frac{3}{4} \end{array}\right) .$$

Now the last three columns are the β-coordinates of the three vectors of α. So

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

A second method to solve this problem uses the standard basis ε and is somewhat more transparant. We know the coordinates w.r.t. the standard basis ε of both the basis α and the basis β, so without computation we know the transition matrices ${}_\varepsilon S_\alpha$ and ${}_\varepsilon S_\beta$:

$${}_\varepsilon S_\alpha = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 2 & 0 & 1 \end{pmatrix} , \quad {}_\varepsilon S_\beta = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix} .$$

For the α to β transition matrix we now have ${}_\beta S_\alpha = {}_\beta S_\varepsilon \; {}_\varepsilon S_\alpha = {}_\varepsilon S_\beta^{-1} \; {}_\varepsilon S_\alpha$.

So we first determine the inverse of ${}_\varepsilon S_\beta$ and find

$${}_\beta S_\varepsilon = \frac{1}{4}\begin{pmatrix} -2 & 2 & 2 \\ 1 & 1 & -1 \\ 3 & -1 & 1 \end{pmatrix} ,$$

and now

$${}_\beta S_\alpha = \frac{1}{4}\begin{pmatrix} -2 & 2 & 2 \\ 1 & 1 & -1 \\ 3 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 2 & 0 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 2 & 4 & -2 \\ -1 & 0 & -3 \\ 5 & -4 & 3 \end{pmatrix} .$$

1.3.13 (Matrix of a linear map) We now discuss how to describe a linear map between two finite dimensional vector spaces using a matrix.

Let V and W be two finite dimensional vector spaces over $\mathbb{K}$ and consider a linear map $\mathcal{A} : V \to W$. After choosing a basis α for V and a basis β for W we are able to represent the linear map $\mathcal{A}$ with a matrix. Consider the following diagram (1.2):

$$\begin{array}{ccc} V & \xrightarrow{\mathcal{A}} & W \\ {\scriptstyle\alpha}\big\downarrow & & \big\downarrow{\scriptstyle\beta} \\ \mathbb{K}^n & \xrightarrow{\beta\mathcal{A}\alpha^{-1}} & \mathbb{K}^m \end{array} \tag{1.2}$$

To every vector $\underline{x} \in V$ there corresponds a unique sequence of coordinates $\alpha(\underline{x})$ and to the image vector $\mathcal{A}\underline{x}$ a sequence $\beta(\mathcal{A}\underline{x})$. The composite linear map $\beta\mathcal{A}\alpha^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ associates to $\alpha(\underline{x})$ the sequence $\beta(\mathcal{A}\underline{x})$. This is a linear map from $\mathbb{K}^n$ to $\mathbb{K}^m$ and is according to 1.2.3 determined by a matrix ${}_\beta A_\alpha$ whose columns are

$$(\beta\mathcal{A}\alpha^{-1})(\underline{e}_i) = \beta(\mathcal{A}\underline{a}_i)\ , \qquad i = 1, \ldots, n\ .$$

(The i–th column consists of the β–coordinates of the image $\mathcal{A}\underline{a}_i$.) The matrix ${}_\beta A_\alpha$ is called *the matrix of $\mathcal{A}$ w.r.t. the bases α and β*. With the matrix ${}_\beta A_\alpha$ we have a description of the linear map $\mathcal{A}$ on coordinate level: to find for example the image of a vector $\underline{a} \in V$, we determine the coordinate vector $\alpha(\underline{a})$ of $\underline{a}$ and multiply this with the matrix ${}_\beta A_\alpha$; in this way we obtain the coordinate vector of $\mathcal{A}\underline{a}$; what remains is to translate back this coordinate vector to the corresponding vector in W.

If $V = W$ and $\beta = \alpha$, then the corresponding matrix is denoted by A_α; we say that this is *the matrix of $\mathcal{A}$ w.r.t. basis α*. In this course most of our attention will go to this situation.

1.3.14 (Connection with *the* matrix) If $\mathcal{A} : V = \mathbb{K}^n \to W = \mathbb{K}^m$ is a linear map, and α and β are the standard bases for these spaces, then the matrix of $\mathcal{A}$ w.r.t. these bases is just *the* matrix of $\mathcal{A}$ as defined in the previous section. In this case the coordinate maps both are the identity maps.

1.3.15 Example. The matrix of a linear map allows us to see what the image of a vector is. If for instance V is two-dimensional with basis $\alpha = (\underline{a}, \underline{b})$ and the linear map $\mathcal{A} : V \to V$ has matrix

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

w.r.t. α, then the matrix tells us that $\mathcal{A}\underline{a} = 1 \cdot \underline{a} - 2 \cdot \underline{b}$ (first column) and $\mathcal{A}\underline{b} = 4 \cdot \underline{a} + 3 \cdot \underline{b}$ (second column). To compute the image of $\lambda\underline{a} + \mu\underline{b}$ we multiply A_α with $(\lambda, \mu)^\top$:

$$\begin{pmatrix} 1 & 4 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \lambda + 4\mu \\ -2\lambda + 3\mu \end{pmatrix}$$

with the result that the image is $(\lambda + 4\mu)\underline{a} + (-2\lambda + 3\mu)\underline{b}$.

1.3.16 Example. Consider the vector space V of real polynomials of degree at most 2 and the linear map $\mathcal{D} : V \to V$ defined by $\mathcal{D}p = p'$. Take for V the basis $\alpha = \{1, x, x^2\}$. The derivatives of the basis vectors are:

$$\begin{aligned} 0 &= \; 0{\cdot}1 + 0{\cdot}x + 0{\cdot}x^2 \;, \\ 1 &= \; 1{\cdot}1 + 0{\cdot}x + 0{\cdot}x^2 \;, \\ 2x &= \; 0{\cdot}1 + 2{\cdot}x + 0{\cdot}x^2 \;. \end{aligned}$$

The matrix D_α is therefore

$$D_\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

To illustrate this we take a polynomial $p(x) = 2x^2 - 3x + 5$. The coordinate-vector of p w.r.t. α is $(5, -3, 2)$ and

$$D_\alpha \begin{pmatrix} 5 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 4 \\ 0 \end{pmatrix}.$$

$(-3, 4, 0)$ is the coordinate vector of $4x - 3$ and this is indeed the derivative of p.

1.3.17 Example. We work in $\mathbb{R}^2$ with the standard inner product. We determine the matrix P_ε w.r.t. the standard basis ε of the orthogonal projection $\mathcal{P}$ on the line ℓ with equation $2x - 3y = 0$. To determine the coordinates of $\mathcal{P}\underline{e}_1$ and $\mathcal{P}\underline{e}_2$ directly is a bit awkward. We consider first a basis $\alpha = \{\underline{a}_1, \underline{a}_2\}$ with respect to which $\mathcal{P}$ has an easy description. We take $\underline{a}_1 = (3, 2) \in \ell$ and $\underline{a}_2 = (2, -3) \perp \ell$. Then $\mathcal{P}\underline{a}_1 = \underline{a}_1 = 1 \cdot \underline{a}_1 + 0 \cdot \underline{a}_2$ and $\mathcal{P}\underline{a}_2 = \underline{0} = 0 \cdot \underline{a}_1 + 0 \cdot \underline{a}_2$ so that the matrix P_α becomes:

$$P_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} .$$

The transition matrices are

$$ {}_{\varepsilon}S_{\alpha} = \begin{pmatrix} 3 & 2 \\ 2 & -3 \end{pmatrix} \quad \text{and} \quad {}_{\alpha}S_{\varepsilon} = {}_{\varepsilon}S_{\alpha}^{-1} = \frac{1}{13}\begin{pmatrix} 3 & 2 \\ 2 & -3 \end{pmatrix}, $$

giving

$$ P_{\varepsilon} = {}_{\varepsilon}S_{\alpha}\ P_{\alpha}\ {}_{\alpha}S_{\varepsilon} = \frac{1}{13}\begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}. $$

This result can also be obtained with the technique of 1.2.5. The data $\mathcal{P}\underline{a}_1 = \underline{a}_1$ and $\mathcal{P}\underline{a}_2 = \underline{0}$ are put in ε-coordinates in the rows of a matrix

$$ \left(\begin{array}{cc|cc} 3 & 2 & 3 & 2 \\ 2 & -3 & 0 & 0 \end{array}\right), $$

which after row reduction gives

$$ \left(\begin{array}{cc|cc} 1 & 0 & \frac{9}{13} & \frac{6}{13} \\ 0 & 1 & \frac{6}{13} & \frac{4}{13} \end{array}\right) $$

and then

$$ P_{\varepsilon} = \frac{1}{13}\begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}. $$

1.3.18 Our rules for coordinate transformations allow us to change to another basis. This gives us a trivial, but important property:

1.3.19 Theorem. (Effect of change of basis) *Choose in a finite dimensional space V two bases α and β, and suppose $\mathcal{A} : V \to V$ is linear. Then*

$$ A_{\beta} = {}_{\beta}S_{\alpha}\, A_{\alpha}\, {}_{\alpha}S_{\beta}\ . $$

Proof. Consider the following equality between compositions

$$ \beta\mathcal{A}\beta^{-1} = (\beta\alpha^{-1})(\alpha\mathcal{A}\alpha^{-1})(\alpha\beta^{-1}) $$

(by replacing the brackets one sees that on the right hand side the factors α^{-1} and α cancel. Now change to matrices and use 1.2.10. □

1.3.20 In words: The matrix A_{β} transforms the β–coordinates of a vector $\underline{x}$ into the β–coordinates of $\mathcal{A}\underline{x}$. We can do this by first transforming the β–coordinates of $\underline{x}$ into its α–coordinates, then compute the α–coordinates of $\mathcal{A}\underline{x}$ with the matrix A_{α} and finally transform them back to β–coordinates.

1.3.21 Example. Back to example 1.3.16. We now take a different basis, $\beta = \{x^2 - x, x^2 + 3, x^2 - 1\}$ for V and try to find the matrix D_β of $\mathcal{D}$. It is a bit awkward to do this directly, but we may use:

$$D_\beta = {}_\beta S_\alpha \, D_\alpha \, {}_\alpha S_\beta \ .$$

Now

$${}_\alpha S_\beta = \begin{pmatrix} 0 & 3 & -1 \\ -1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} ,$$

so

$${}_\beta S_\alpha = {}_\alpha S_\beta^{-1} = \frac{1}{4} \begin{pmatrix} 0 & -4 & 0 \\ 1 & 1 & 1 \\ -1 & 3 & 3 \end{pmatrix}$$

and

$$D_\beta = {}_\beta S_\alpha \, D_\alpha \, {}_\alpha S_\beta = \frac{1}{4} \begin{pmatrix} -8 & -8 & -8 \\ 1 & 2 & 2 \\ 7 & 6 & 6 \end{pmatrix} .$$

Again as an illustration: $p(x) = 2x^2 - 3x + 5 = 3(x^2 - x) + (x^2 + 3) - 2(x^2 - 1)$ has β–coordinates $(3, 1, -2)$,

$$D_\beta \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -8 & -8 & -8 \\ 1 & 2 & 2 \\ 7 & 6 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -16 \\ 1 \\ 15 \end{pmatrix} .$$

These now should be the β–coordinates of the derivative of $p(x)$ which is correct, since $-4(x^2 - x) + \frac{1}{4}(x^2 + 3) + \frac{15}{4}(x^2 + 1) = 4x - 3 = p'(x)$.

1.4 Eigenvalues and eigenvectors, diagonalisation

1.4.1 Let $\mathcal{A} : V \to V$ be a linear map with V finite dimensional. For every choice of a basis α for V the map $\mathcal{A}$ is determined by a matrix A_α ; if α and β are different bases then the connection between A_α and A_β is given by Theorem 1.3.19. In the two examples we have already seen that for a suitably chosen basis α the matrix A_α may have a pleasant simple form. It turns out that such simple forms can be found systematically. In this section we discuss the relevant techniques. First we make more precise what we mean by "simple form".

1.4.2 Definition. A square matrix A has a *diagonal form* (or shorter, is diagonal) if all elements a_{ij} with $i \neq j$ are equal to zero.

The following theorem is easy to understand, but of fundamental importance:

1.4.3 Theorem. *Let $\mathcal{A} : V \to V$ be a linear map and $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ be a basis for V. The matrix A_α has the diagonal form*

$$A_\alpha = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

if and only if $\mathcal{A}\underline{a}_i = \lambda_i \underline{a}_i$ for $i = 1, \ldots, n$.

Proof. The matrix A_α has the diagonal form above if and only if for every i the α–coordinates of $\mathcal{A}\underline{a}_i$ are equal to $(0, \ldots, 0, \lambda_i, 0, \ldots, 0)$, so if and only if $\mathcal{A}\underline{a}_i = 0{\cdot}\underline{a}_1 + \cdots + \lambda_i{\cdot}\underline{a}_i + \cdots + 0{\cdot}\underline{a}_n = \lambda_i \underline{a}_i$. □

1.4.4 Definition. (Eigenvector and eigenvalue) Let $\mathcal{A} : V \to V$ be a linear map. A vector $\underline{v} \neq \underline{0}$ is called *eigenvector* of $\mathcal{A}$ with *eigenvalue* λ if $\mathcal{A}\underline{v} = \lambda\underline{v}$.

So eigenvectors are vectors *different from the zero vector* that are mapped by $\mathcal{A}$ on a multiple of itself; this multiple is the corresponding eigenvalue.

Theorem 1.4.3 can therefore also be formulated as follows:

1.4.5 Theorem. *Let $\mathcal{A} : V \to V$ be a linear map. The matrix A_α is in diagonal form if and only if α is a basis of eigenvectors. In that case the diagonal elements are the eigenvalues.*

1.4.6 Example. In example 1.3.17 we have looked at the projection on a line. The vectors $\underline{a}_1$ and $\underline{a}_2$ we defined there are eigenvectors with eigenvalues 1 and 0 respectively; the matrix w.r.t. the basis $\{\underline{a}_1, \underline{a}_2\}$ is therefore

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

a diagonal matrix with the eigenvalues on the diagonal. Notice the order of the 1 and the 0 on the diagonal: this corresponds to the order of the eigenvectors.

1.4.7 Example. Consider in the euclidean plane E^2 a rotation over 90° around the origin. No vector different from $\underline{0}$ is mapped to a multiple of itself. So this linear map has no eigenvectors and certainly no basis of eigenvectors. There is no choice of basis for which the matrix of this rotation is diagonal.

1.4.8 Definition. (Eigenspace) Consider a linear map $\mathcal{A} : V \to V$. For every number λ let

$$E_\lambda = \mathcal{N}(\mathcal{A} - \lambda\mathcal{I}) \ ;$$

E_λ is called the *eigenspace* of $\mathcal{A}$ for λ.

1.4.9 (Finding eigenvalues and eigenvectors) E_λ is the null space of the linear map $\mathcal{A} - \lambda\mathcal{I}$. We have $\underline{v} \in E_\lambda$ if and only if $(\mathcal{A} - \lambda\mathcal{I})\underline{v} = \underline{0}$, so if and only if $\mathcal{A}v - \lambda\underline{v} = \underline{0}$, that is, if and only if $\mathcal{A}\underline{v} = \lambda\underline{v}$. The eigenspace E_λ consists therefore of the zero vector together with all eigenvectors for eigenvalue λ. A special eigenspace is E_0. This consists of vectors that are mapped to 0 times itself so on $\underline{0}$. So E_0 is the null space of $\mathcal{A}$.

As a rule E_λ will just contain $\{\underline{0}\}$. We say that λ is an eigenvalue if and only if E_λ constains a vector $\underline{v} \neq \underline{0}$, so if and only if $\dim(E_\lambda) > 0$. Some authors only use the word eigenspace if λ is an eigenvalue.

The procedure we follow in the search for eigenvectors is to first determine λ for which $\dim(E_\lambda) > 0$. Then λ is an eigenvalue and the vectors in E_λ different from $\underline{0}$ are the eigenvectors for this eigenvalue.

Choose any basis $\alpha = \{\underline{a}_1, \dots, \underline{a}_n\}$ for V (so not necessarily a basis of eigenvectors). Let A be the matrix of $\mathcal{A}$ w.r.t. this basis. Then

$$A - \lambda I = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} - \lambda \end{pmatrix}$$

is the matrix of $\mathcal{A} - \lambda\mathcal{I}$ w.r.t. this basis α .

Let $\underline{v} = v_1\underline{a}_1 + \dots + v_n\underline{a}_n$ be a vector in V. Then $\underline{v}$ is eigenvector for eigenvalue λ if and only if $\underline{v} \neq \underline{0}$ and $(\mathcal{A} - \lambda\mathcal{I})\underline{v} = \underline{0}$, so if and only if $(v_1, \dots, v_n) \neq (0, \dots, 0)$ and

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} . \tag{1.3}$$

So we have eigenvectors for λ if and only the dimension of the solution space of the homogeneous system $(A - \lambda I)\underline{x} = \underline{0}$ is bigger than 0, and that is the case if and only if the rank of $A - \lambda I$ is less than n, and that happens if and only if $\det(A - \lambda I) = 0$. If λ is an eigenvalue, dan the solutions of the system $(A - \lambda I)\underline{x} = \underline{0}$ are the coordinate vectors of elements of E_λ.

1.4.10 Theorem. *λ is an eigenvalue if and only if* $\det(A - \lambda I) = 0$ *and the α-coordinates of the corresponding eigenvectors are the solutions different from the zero solution of the system (1.3).*

1.4.11 Definition. (Characteristic polynomial) Let $\mathcal{A} : V \to V$ be a linear map and let A_α be the matrix of $\mathcal{A}$ w.r.t. a basis α. Then the equation $\det(A_\alpha - \lambda I) = 0$ is called the *characteristic equation* of $\mathcal{A}$. The left hand side of this equation, $\det(A_\alpha - \lambda I)$, is called the *characteristic polynomial* of $\mathcal{A}$ (and of the matrix A_α).

1.4.12 Since eigenvalues and eigenvectors only depend on $\mathcal{A}$ and not on the choice of a particular basis α the characteristic equation should not depend on this basis. To show that this is indeed the case we consider what happens if we take a different basis β.

$$\begin{aligned} \det(A_\beta - \lambda I) &= \det({}_\beta S_\alpha A_{\alpha\alpha} S_\beta - \lambda_\beta S_\alpha I_\alpha S_\beta) = \\ = \det({}_\beta S_\alpha (A_\alpha - \lambda I)_\alpha S_\beta) &= \det({}_\beta S_\alpha) \det(A_\alpha - \lambda I) \det({}_\alpha S_\beta) = \\ = \det(A_\alpha - \lambda I) \det({}_\beta S_{\alpha\alpha} S_\beta) &= \det(A_\alpha - \lambda I) \det(I) = \\ &= \det(A_\alpha - \lambda I) \ . \end{aligned}$$

We see that the matrices A_α and A_β have the same characteristic polynomial.

1.4.13 Before discussing examples we examine the characteristic polynomial in more detail. This polynomial is the determinant

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} - \lambda \end{vmatrix} .$$

This determinant is the sum of $n!$ terms and every term is a product of n matrix elements, one from each row and one from each column. Every term in this sum is therefore a polynomial in λ of degree at most n. One of the terms is the product of the elements on the main diagonal:

$$\begin{aligned} &(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) \\ &\quad = (-1)^n \lambda^n + (-1)^{n-1} (a_{11} + a_{22} + \cdots + a_{nn}) \lambda^{n-1} + \cdots . \end{aligned}$$

Of the remaining terms each one contains an element a_{ij} with $i \neq j$. That means that in such a term the diagonal elements $(a_{ii} - \lambda)$ and $(a_{jj} - \lambda)$ do not occur, so what is left is a polynomial of degree at most $n-2$ in λ. The characteristic polynomial therefore is a polynomial of degree (exactly) n of the following shape

$$(-1)^n\lambda^n + (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + c_1\lambda + c_0 \ .$$

1.4.14 Definition. (Trace) The sum of the diagonal elements of a square matrix is called the *trace* of the matrix.

Because the characteristic equation does not depend on the choice of a basis we see:

1.4.15 Theorem. *Let $\mathcal{A} : V \to V$ be a linear map with* $\dim(V) < \infty$. *For every basis α the matrix A_α has the same trace.*

The coefficient c_0 of the characteristic polynomial is also easy to find. Taking $\lambda = 0$ we see $\det(A) = c_0$. So:

1.4.16 Theorem. *Let $\mathcal{A} : V \to V$ be a linear map with* $\dim(V) < \infty$. *For every basis α the matrix A_α has the same determinant.*

1.4.17 Let $\lambda_1, \ldots, \lambda_n$ be the n roots (in $\mathbb{C}$, with multiplicity) of the characteristic equation. Then the characteristic polynomial can be written as

$$\begin{aligned}&(-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n) = \\ &\qquad = (-1)^n\lambda^n + (-1)^{n-1}(\lambda_1 + \lambda_2 + \ldots + \lambda_n)\lambda^{n-1} + \cdots + \lambda_1\lambda_2\cdots\lambda_n.\end{aligned}$$

As a consequence we have:

1.4.18 Theorem. *The trace of the matrix is the sum of the roots of the characteristic equation and the determinant of the matrix is the product of the roots of the characteristic equation.*

1.4.19 Eigenvalues and eigenvectors can therefore be determined as follows (although there may sometimes be more clever ways):

- Determine the matrix A_α of the linear map $\mathcal{A} : V \to V$ w.r.t. some basis α.
- Compute the characteristic equation $\det(A_\alpha - \lambda I) = 0$ and solve it. In a real vector space, only the real roots are eigenvalues of the linear map. In a complex space, every root is an eigenvalue.

- Solve for every eigenvalue λ the system $(A_\alpha - \lambda I)\underline{v} = \underline{0}$ (see 1.3). Every solution gives the *coordinates* of a vector from E_λ. The solutions form the eigenspace E_λ, in terms of coordinates w.r.t. the basis α.
- If necessary transform back from coordinates to vectors in V.

In case there is a basis of eigenvectors, then forming the matrix w.r.t. such a basis is simple: the matrix is a diagonal matrix and on the diagonal are the eigenvalues (in the same order as the corresponding eigenvectors in the basis). So there is no need to compute explicit transformations.

1.4.20 Example. We consider in $\mathbb{R}^2$ with the standard inner product the projection on the line $\ell : 2x - 3y = 0$ (see example 1.3.17). The matrix of this projection is

$$\frac{1}{13}\begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix},$$

the characteristic equation therefore is

$$\begin{vmatrix} \frac{9}{13} - \lambda & \frac{6}{13} \\ \frac{6}{13} & \frac{4}{13} - \lambda \end{vmatrix} = \frac{1}{169}\begin{vmatrix} 9 - 13\lambda & 6 \\ 6 & 4 - 13\lambda \end{vmatrix}$$

$$= \frac{1}{169}\Big((9 - 13\lambda)(4 - 13\lambda) - 36\Big) = \lambda^2 - \lambda = 0 .$$

The characteristic equation has two roots: $\lambda = 1$ and $\lambda = 0$. The trace of the matrix is 1 and the determinant is 0.

The eigenspace for $\lambda = 1$ can be found from the system of equations with coefficient matrix

$$\begin{pmatrix} \frac{9}{13} - 1 & \frac{6}{13} \\ \frac{6}{13} & \frac{4}{13} - 1 \end{pmatrix} \sim \begin{pmatrix} -4 & 6 \\ 6 & -9 \end{pmatrix} \sim \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix}$$

so from $2x - 3y = 0$ (the equation of ℓ, no surprise). All solutions (x, y) are therefore multiples of $(3, 2)$, so

$$E_1 =< 3\underline{e}_1 + 2\underline{e}_2 > (= \ell) .$$

The eigenspace for $\lambda = 0$ can be found from the system

$$\begin{pmatrix} \frac{9}{13} - 0 & \frac{6}{13} \\ \frac{6}{13} & \frac{4}{13} - 0 \end{pmatrix} \sim \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix} \sim \begin{pmatrix} 3 & 2 \\ 0 & 0 \end{pmatrix},$$

that is $3x + 2y = 0$. Hence all solutions (x, y) are multiples of $(2, -3)$, so

$$E_0 = < 2\underline{e}_1 - 3\underline{e}_2 > (= \ell^{\perp}) .$$

Take $\underline{a}_1 = 3\underline{e}_1 + 2\underline{e}_2$ and $\underline{a}_2 = 2\underline{e}_1 - 3\underline{e}_2$. Then $\alpha = \{\underline{a}_1, \underline{a}_2\}$ is a basis of eigenvectors and the matrix of the projection w.r.t. this basis has the diagonal form

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

Just as the original matrix, this matrix has trace 1 and determinant 0 in accordance with Theorems 1.4.15 and 1.4.16.

1.4.21 Example. We choose in the euclidean plane E^2 an orthonormal basis $\{\underline{e}_1, \underline{e}_2\}$ and consider a rotation over $90°$ sending $\underline{e}_1$ to $\underline{e}_2$. So $\mathcal{A}\underline{e}_1 = \underline{e}_2$ and $\mathcal{A}\underline{e}_2 = -\underline{e}_1$ and the the matrix w.r.t. $\{\underline{e}_1, \underline{e}_2\}$ becomes

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

The characteristic equation is

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0$$

and has two roots: $\lambda = i$ and $\lambda = -i$. The trace is 0 and the determinant is 1. Since E^2 is a real vector space i and $-i$ cannot be eigenvalues; so this rotation has no eigenvectors, in agreement example 1.4.7.

1.4.22 Example. Consider the vector space V of real polynomials of degree at most two and define the linear map $\mathcal{A} : V \to V$ by $\mathcal{A}p = p'$. Then w.r.t. the basis $\{1, x, x^2\}$, $\mathcal{A}$ has the matrix (see example 1.3.16)

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

The characteristic equation is therefore

$$\begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda^3 = 0 .$$

The only root of the characteristic equation is $\lambda = 0$ with multiplicity 3; the trace is 0 and the determinant is 0. We find the eigenspace E_0 by solving the homogeneous system with coefficient matrix

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

All solutions of this system of equations are multiples of $(1, 0, 0)$, so

$$E_0 =< 1{\cdot}1 + 0{\cdot}x + 0{\cdot}x^2 >=< 1 > .$$

E_0 therefore consists of the constant polynomials. They indeed have the zero polynomial as derivative. Notice that $\dim(E_0) = 1$ although $\lambda = 0$ has multiplicity 3. So we do not find a basis of eigenvectors, and also no diagonal form for the matrix of the map.

1.4.23 Example. The linear map $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

$$A = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix} .$$

We work w.r.t. the standard basis $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ which means that in this example we may identify vectors and coordinates. The characteristic equation is:

$$\begin{vmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{vmatrix} = 0, \quad \text{so} \quad -\lambda^3 + 13\lambda^2 - 40\lambda + 36 = 0.$$

This polyomial factors as $-(\lambda - 9)(\lambda - 2)^2$; so there are two eigenvalues, $\lambda = 9$ and $\lambda = 2$, the last one with multiplicity 2. The trace is indeed $9 + 2 + 2 = 13$ and the determinant is 36. Check this.
We find the eigenspace E_9 from the system of equations with matrix

$$\begin{pmatrix} -5 & -1 & 6 \\ 2 & -8 & 6 \\ 2 & -1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

with solutions the multiples of $(1, 1, 1)$. So $E_9 =< (1, 1, 1) >$.
The eigenspace E_2 we find from the equations

$$\begin{pmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{pmatrix},$$

so E_2 is the plane $2x - y + 6z = 0$: $E_2 =< (1,2,0),(0,6,1) >$. The basis $\alpha = \{(1,1,1),(1,2,0),(0,6,1)\}$ is a basis of eigenvectors. W.r.t. this basis

$$A_\alpha = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

A_α is a diagonal matrix, of course with trace 13 and determinant 36.

1.4.24 From the examples above we see that for a given linear map $\mathcal{A} : V \to V$ we can not always find a basis of eigenvectors, which implies that we can not always represent $\mathcal{A}$ by a diagonal matrix. Still it is often possible to find a reasonably simple matrix by chosing a suitable basis. This will be the subject of the next section.

To determine whether a system of eigenvectors is independent, the following result is very useful:

1.4.25 Theorem. *Let $\mathcal{A} : V \to V$ be a linear map and let $\underline{v}_1, \ldots, \underline{v}_n$ be eigenvectors of $\mathcal{A}$ for* mutually different *eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $\underline{v}_1, \ldots, \underline{v}_n$ are independent.*

Proof. Suppose not. We may assume that $\underline{v}_1$ is dependent of $\underline{v}_2, \ldots, \underline{v}_n$; if necessary we renumber the vectors. Next we prune $\underline{v}_2, \ldots, \underline{v}_n$ to an independent system $\underline{v}_2, \ldots, \underline{v}_p$, so $< \underline{v}_2, \ldots, \underline{v}_p > = < \underline{v}_2, \ldots, \underline{v}_n >$; to accomplish this we might have to renumber the vectors again. We now have

$$\underline{v}_1 = \sum_{i=2}^{p} \alpha_i \underline{v}_i \quad \text{implying} \quad \lambda_1 \underline{v}_1 = \sum_{i=2}^{p} \alpha_i \lambda_1 \underline{v}_i = \mathcal{A}\underline{v}_1 = \mathcal{A} \sum_{i=2}^{p} \alpha_i \underline{v}_i = \sum_{i=2}^{p} \alpha_i \lambda_i \underline{v}_i,$$

and so

$$\sum_{i=2}^{p} \alpha_i (\lambda_i - \lambda_1) \underline{v}_i = \underline{0}.$$

Since $\underline{v}_2, \ldots, \underline{v}_p$ are independent and $\lambda_i - \lambda_1 \neq 0$ for $i = 2, \ldots, p$, it follows that $\alpha_i = 0$ for $i = 2, \ldots, p$, and so $\underline{v}_1 = \underline{0}$ which is impossible because $\underline{v}_1$ is an eigenvector. The system $\underline{v}_1 \ldots, \underline{v}_n$ is therefore independent. □

1.4.26 A final remark. In this section we have been looking for methods to represent a linear map by a diagonal matrix; the starting point in this case is a linear map. We can also start from a square matrix A, view this as the matrix of a linear map $\mathbb{K}^n \to \mathbb{K}^n$ and look for a basis α of eigenvectors (provided such a basis exists). Collect the eigenvalues in a diagonal matrix D. This is now a diagonal form of the matrix A. So we have $D = {}_\alpha S_\varepsilon \, A \, {}_\varepsilon S_\alpha$, equivalently $A = {}_\varepsilon S_\alpha \, D \, {}_\alpha S_\varepsilon$. This procedure is called *to diagonalise* the matrix A.

1.5 Invariant subspaces

1.5.1 Eigenvectors have the property that they are mapped to a multiple of itself; the line spanned by the vector is then mapped into itself. If there are no, or not enough eigenvectors, there still might be subspaces that are mapped to itself, so called invariant subspaces. Imagine for instance a rotation about an axis in $\mathbb{R}^3$, in this case the plane perpendicular to the axis is such an invariant subspace (see Chapter 2). Such invariant subspaces also lead to simple matrix forms. In this section we will discuss

- the notion invariant subspace of a linear map,
- the role of invariant subspaces in the determination of simple forms for a matrix,
- a technique to determine an invariant subspace, by using a non–real root of the characteristic equation of a linear map of a real vector space to itself.

All maps considered in this section are $\mathcal{A} : V \to V$ with $\dim(V) < \infty$.

1.5.2 Definition. (Invariant subspace) Let W be a subspace of V. W is called *invariant* under the linear map $\mathcal{A} : V \to V$ if $\mathcal{A}\underline{w} \in W$ for all $\underline{w} \in W$.

1.5.3 Example. The null space $\mathcal{N}$ of a linear map $\mathcal{A}$ is invariant: if $\underline{x} \in \mathcal{N}$, then $\mathcal{A}\underline{x} = \underline{0}$ and this vector again belongs to $\mathcal{N}$.

The range $\mathcal{R}$ is also invariant: if $\underline{y} \in \mathcal{R}$, then $\mathcal{A}\underline{y}$ is obviously again contained in $\mathcal{R}$.

1.5.4 If W is invariant under $\mathcal{A}$, then all image vectors $\mathcal{A}\underline{w}$ with $\underline{w} \in W$ are again in W. So if we restrict $\mathcal{A}$ to W we have obtained a linear map $\mathcal{A} : W \to W$, the restriction of $\mathcal{A}$ tot W.

Fortunately we usually don't have to check that *every* single vector of W is mapped by $\mathcal{A}$ into W:

1.5.5 Theorem. *Let $\mathcal{A} : V \to V$ be linear and let $W =< \underline{a}_1, \ldots, \underline{a}_n >$. W is invariant under $\mathcal{A}$ if and only if $\mathcal{A}\underline{a}_i \in< \underline{a}_1, \ldots, \underline{a}_n >$ for $i = 1, \ldots, n$.*

Proof. If W is invariant, then $\mathcal{A}\underline{w} \in W$ for all $\underline{w} \in W$, so in particular for $\underline{w} = \underline{a}_1, \ldots, \underline{a}_n$. Conversely, suppose $\mathcal{A}\underline{a}_i \in W$ for $i = 1, \ldots, n$ and take an arbitrary $\underline{w} \in W$. We may write $\underline{w}$ as $\underline{w} = w_1\underline{a}_1 + \cdots + w_n\underline{a}_n$, and now $\mathcal{A}\underline{w} = w_1\mathcal{A}\underline{a}_1 + \cdots + w_n\mathcal{A}\underline{a}_n$. Since every $\mathcal{A}\underline{a}_i \in W$, this implies that $\mathcal{A}\underline{w} \in W$. $\square$

1.5.6 Theorem. *Suppose $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ is a basis for V such that $W =< \underline{a}_1, \ldots, \underline{a}_m >$ is invariant under $\mathcal{A}$. Then the matrix A_α has the following form*

$$A_\alpha = \begin{pmatrix} & & & * & \ldots & * \\ & M_1 & & \vdots & & \vdots \\ 0 & \ldots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & * & \ldots & * \end{pmatrix},$$

The $m \times m$–matrix M_1 is the matrix of the restriction $\mathcal{A} : W \to W$ w.r.t. the basis $\{\underline{a}_1, \ldots, \underline{a}_m\}$.

Proof. For $i = 1, \ldots, m$ the previous theorem gives that

$$\mathcal{A}\underline{a}_i = a_{i1}\underline{a}_1 + \cdots + a_{im}\underline{a}_m + 0\underline{a}_{m+1} + \cdots + 0\underline{a}_n \ .$$

From this the claim follows immediately □

1.5.7 Example. Consider in $\mathbb{R}^5$ the (independent) vectors

$$\underline{a} = (1, 0, -2, 2, 1) \text{ and } \underline{b} = (2, 0, 2, -1, -1).$$

The linear map $\mathcal{A} : \mathbb{R}^5 \to \mathbb{R}^5$ has the property that

$$\mathcal{A}\underline{a} = (0, 0, -6, 5, 3)\ ,\ \mathcal{A}\underline{b} = (5, 0, 2, 0, -1).$$

We are going to show that $W =< \underline{a}, \underline{b} >$ is invariant and we will determine a matrix of the restriction $\mathcal{A} : W \to W$.

To show the invariance of $< \underline{a}, \underline{b} >$, we must verify that $\mathcal{A}\underline{a}$ and $\mathcal{A}\underline{b}$ are linear combinations of $\underline{a}$ and $\underline{b}$. This we do by simultaneously solve the systems of equations with columns $\underline{a}, \underline{b}, \mathcal{A}\underline{a}, \mathcal{A}\underline{b}$:

$$\left(\begin{array}{rr|rr} 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ -2 & 2 & -6 & 2 \\ 2 & -1 & 5 & 0 \\ 1 & -1 & 3 & -1 \end{array}\right).$$

After row reduction and deleting zero rows we find

$$\left(\begin{array}{rr|rr} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 2 \end{array}\right),$$

and it follows that $\mathcal{A}\underline{a} = 2\underline{a} - \underline{b}$ and $\mathcal{A}\underline{b} = \underline{a} + 2\underline{b}$. So $W =< \underline{a}, \underline{b} >$ is invariant under $\mathcal{A}$, and the matrix of $\mathcal{A} : W \to W$ w.r.t. the basis $\{\underline{a}, \underline{b}\}$ is

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

1.5.8 (Two-dimensional invariant subspaces) For non-real roots of the characteristic equation of a linear map $\mathcal{A} : V \to V$ of a *real* vector space V to itself we have no corresponding eigenvectors. But such a root does give us a two-dimensional invariant subspace. This we will demonstrate now.

Choose a basis $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ for V and consider the characteristic equation $\det(A_\alpha - \lambda I) = 0$. Let μ be a non-real root of this equation. Then μ is not an eigenvalue of $\mathcal{A} : V \to V$, but considering A_α as the matrix of an auxiliary map $\mathbb{C}^n \to \mathbb{C}^n$ then μ is an eigenvalue and we find a complex vector $\underline{z} \neq \underline{0}$ such that

$$A_\alpha \underline{z} = \mu \underline{z} \; .$$

Suppose $(x_1, \ldots, x_n) = (\mathrm{Re} z_1, \ldots, \mathrm{Re} z_n)$ (so '$\underline{x} = \mathrm{Re}\underline{z}$') and $(y_1, \ldots, y_n) = (\mathrm{Im} z_1, \ldots, \mathrm{Im} z_n)$. Then $\underline{z} = \underline{x} + i\,\underline{y}$ and the rules for matrix multiplication give

$$A_\alpha \underline{z} = A_\alpha(\underline{x} + i\,\underline{y}) = A_\alpha \underline{x} + i\, A_\alpha \underline{y}$$

and

$$\mu\; \underline{z} = \mathrm{Re}(\mu)\; \underline{x} - \; \mathrm{Im}(\mu)\; \underline{y} + i\; \mathrm{Im}(\mu)\; \underline{x} + i\; \mathrm{Re}(\mu)\; \underline{y}.$$

This implies that

$$A_\alpha \underline{x} = \; \mathrm{Re}(\mu)\; \underline{x} - \; \mathrm{Im}(\mu)\; \underline{y}, \qquad A_\alpha \underline{y} = \; \mathrm{Im}(\mu)\; \underline{x} + \; \mathrm{Re}(\mu)\; \underline{y}.$$

Consider now $\underline{x}$ and $\underline{y}$ as vectors in V (where we just write $\underline{x}$ and $\underline{y}$ instead of what really should be $\alpha^{-1}(\underline{x})$ and $\alpha^{-1}\underline{y}$): $\underline{x} = x_1\underline{a}_1 + \cdots + x_n\underline{a}_n \in V$ and $\underline{y} = y_1\underline{a}_1 + \cdots + y_n\underline{a}_n \in V$. Then

$$\begin{aligned} \mathcal{A}\underline{x} &= \mathrm{Re}(\mu)\; \underline{x} \;-\; \mathrm{Im}(\mu)\; \underline{y}, \\ \mathcal{A}\underline{y} &= \mathrm{Im}(\mu)\; \underline{x} \;+\; \mathrm{Re}(\mu)\; \underline{y}. \end{aligned}$$

The space $W =< \underline{x}, \underline{y} >$ is a two-dimensional (check this!) invariant subspace and w.r.t. the basis $\{\underline{x}, \underline{y}\}$, the restriction $\mathcal{A} : W \to W$ has the matrix:

$$\begin{pmatrix} \mathrm{Re}(\mu) & \mathrm{Im}(\mu) \\ -\mathrm{Im}(\mu) & \mathrm{Re}(\mu) \end{pmatrix}$$

with μ and $\overline{\mu}$ as roots of the characteristic equation. We will use this again in the next chapter.

1.5.9 Let W be an invariant subspace for a linear map $\mathcal{A} : V \to V$ and choose a basis $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ for V in such a way that $\{\underline{a}_1, \ldots, \underline{a}_m\}$ is a basis for W. Then Theorem 1.5.6 implies

$$A_\alpha - \lambda I = \begin{pmatrix} & * & \ldots & * \\ M_1 - \lambda I^* & \vdots & & \vdots \\ 0 \ldots 0 & \vdots & & \vdots \\ \vdots\vdots & \vdots & M_2 & \vdots \\ 0 \ldots 0 & * & \ldots & * \end{pmatrix}.$$

Here I^* is the $m \times m$ identity matrix. From this structure we see that

$$\det(A_\alpha - \lambda I) = \det(M_1 - \lambda I^*)\det(M_2).$$

(More generally one can show, using induction and cofactor expansion along the first row that

$$\det\begin{pmatrix} A & * \\ O & B \end{pmatrix} = \det A \, \det B,$$

for a $p \times p$–matrix A, and a $q \times q$–matrix B, while $*$ stands for an arbitrary $p \times q$–matrix O for the $q \times p$–zero matrix.) On the left we have the characteristic polynomial of $\mathcal{A} : V \to V$ while the first factor on the right is the characteristic polynomial of the restriction $\mathcal{A} : W \to W$. We find the following important property:

1.5.10 Theorem. *If W is an invariant subspace for the linear map $\mathcal{A} : V \to V$, then the characteristic polynomial of the restriction $\mathcal{A} : W \to W$ is a factor of the characteristic polynomial of $\mathcal{A} : V \to V$.*

1.5.11 The previous theorem has an important consequence. Consider a linear map $\mathcal{A} : V \to V$. If $\underline{x}$ is an eigenvector with eigenvalue μ, then also $\mathcal{A}\underline{x}$ is an eigenvector for eigenvalue μ (assuming $\mathcal{A}\underline{x} \neq \underline{0}$), check this. It follows that for every eigenvalue μ the eigenspace E_μ is invariant under $\mathcal{A}$. Choose a basis $\{\underline{a}_1, \ldots, \underline{a}_m\}$ for E_μ. Then w.r.t. this basis the $m \times m$ matrix of the restriction $\mathcal{A} : E_\mu \to E_\mu$ is:

$$\begin{pmatrix} \mu & & 0 \\ & \ddots & \\ 0 & & \mu \end{pmatrix} = \mu \, I_m,$$

with characteristic polynomial $(\mu - \lambda)^m$. So $(\mu - \lambda)^m$ is a factor of the characteristic polynomial of $\mathcal{A} : V \to V$, implying

1.5.12 Theorem. *The dimension of the eigenspace for eigenvalue μ of a linear map $\mathcal{A} : V \to V$ is smaller than or equal to the multiplicity of μ as a root of the characteristic equation.*

The following theorem is a slight extension of Theorem 1.5.6. Supply your own proof.

1.5.13 Theorem. *Let $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ be a basis for V such that $W_1 =< \underline{a}_1, \ldots, \underline{a}_m >$ and $W_2 =< \underline{a}_{m+1}, \ldots, \underline{a}_n >$ are invariant under $\mathcal{A} : V \to V$. Then the matrix A_α has the form*

$$A_\alpha = \begin{pmatrix} \begin{matrix} & \\ M_1 & \\ & \end{matrix} & \begin{matrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{matrix} \\ \begin{matrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{matrix} & \begin{matrix} & \\ M_2 & \\ & \end{matrix} \end{pmatrix} .$$

Here M_1 and M_2 are the $m \times m$ and $(n-m) \times (n-m)$ matrices of the restrictions $\mathcal{A} : W_1 \to W_1$ and $\mathcal{A} : W_2 \to W_2$ w.r.t. the bases $\{\underline{a}_1, \ldots, \underline{a}_m\}$ and $\{\underline{a}_{m+1}, \ldots, \underline{a}_n\}$ respectively. In addition we have that $\det(A_\alpha) = \det(M_1)\ \det(M_2)$.

1.5.14 At the end of this section we indicate what can be done if the dimension of the eigenspace for an eigenvalue is smaller than the multiplicity of the eigenvalue as a root of the characteristic equation.

Consider a linear map $\mathcal{A} : V \to V$. Let $\mathcal{R}$ be the range and $\mathcal{N}$ the null space. $\mathcal{R}$ and $\mathcal{N}$ are invariant subspaces under $\mathcal{A}$.

Consider first the situation that $\mathcal{R}$ and $\mathcal{N}$ have only the zero vector in common. Then the null space of the restriction $\mathcal{A} : \mathcal{R} \to \mathcal{R}$ consists of just zero vector; in other words, the map $\mathcal{A} : \mathcal{R} \to \mathcal{R}$ has no eigenvalue 0. Choose a basis $\{\underline{a}_1, \ldots, \underline{a}_p\}$ for $\mathcal{N}$ and a basis $\{\underline{a}_{p+1}, \ldots, \underline{a}_n\}$ for $\mathcal{R}$. This is possible because of dimension theorem. Since $\mathcal{R} \cap \mathcal{N} = \{\underline{0}\}$ $\{\underline{a}_1, \ldots, \underline{a}_p, \underline{a}_{p+1}, \ldots, \underline{a}_n\}$ is a basis for V so the matrix of $\mathcal{A}$ w.r.t. this basis looks like

$$\begin{pmatrix} 0 & & & 0 & \ldots & \ldots & 0 \\ & \ddots & & \vdots & & & \vdots \\ & & 0 & 0 & \ldots & \ldots & 0 \\ 0 & \ldots & 0 & & & & \\ \vdots & & \vdots & & M & & \\ \vdots & & \vdots & & & & \\ 0 & \ldots & 0 & & & & \end{pmatrix} .$$

Here M is the $(n-p)\times(n-p)$ matrix is of the restriction $\mathcal{A}:\mathcal{R}\to\mathcal{R}$. It follows that the characteristic equation of $\mathcal{A}:V\to V$ equals

$$(-\lambda)^p\cdot\det(M-\lambda I^*)=0\ .$$

Since $\mathcal{A}:\mathcal{R}\to\mathcal{R}$ has no eigenvalue $\lambda=0$, $\det(M-\lambda I^*)$ has no factor λ and $\lambda=0$ is therefore a p-fold root of the characteristic equation of $\mathcal{A}$. So: if $\mathcal{R}\cap\mathcal{N}=\{\underline{0}\}$, then the dimension of the null space is equal to the multiplicity of $\lambda=0$ as a root of the characteristic equation.

From this we then conclude that if the multiplicity of $\lambda=0$ as a root of the characteristic equation is larger than the dimension of the null space, then $\mathcal{R}\cap\mathcal{N}$ contains more than just the zero vector; so there exists a vector $\underline{a}$ with $\mathcal{A}\underline{a}\neq\underline{0}$ and $\mathcal{A}^2\underline{a}=\underline{0}$.

When we apply this result to $\mathcal{A}-\mu I$ in stead of $\mathcal{A}$ it yields that if the multiplicity of $\lambda=\mu$ as a root of the characteristic equation is larger than the dimension of the eigenspace E_μ, then there is a vector $\underline{a}$ with $(\mathcal{A}-\mu I)\underline{a}=\underline{b}\neq\underline{0}$ and $(\mathcal{A}-\mu I)\underline{b}=\underline{0}$. This can be used to find *upper triangular forms* for a linear map. We don't elaborate on this.

1.6 Dual spaces

1.6.1 Property 1.1.17 implies that we may add linear maps (from V to W for fixed V and W) and multiply them with scalars. It is easy to check that these operations satisfy the rules for a vector space with the obvious zero vector and opposite. So with these operations the set of linear maps from V to W becomes a vector space. A special and important case occurs if W is the field $\mathbb{K}$ of scalars, (recall that $\mathbb{K}$ stands for $\mathbb{R}$ of $\mathbb{C}$). In this case we call our maps *linear functions* or *linear functionals* on V. The space of linear functions, the *dual space*, is the subject of this section.

1.6.2 Definition. (Dual space) Let V be a vector space over $\mathbb{K}$. The vector space of linear maps $V\to\mathbb{K}$ is called the *dual space of* V. We denote this vector space by V^*.

1.6.3 Example. The "sum of the first and second coordinate" is a linear function ϕ on $\mathbb{R}^3$, so an element of $(\mathbb{R}^3)^*$. In a formula: $\phi(x_1,x_2,x_3)=x_1+x_2$. The function $\mathbb{R}^3\to\mathbb{R}$, $(x_1,x_2,x_3)\mapsto x_1^2$ is not linear, so does not belong to $(\mathbb{R}^3)^*$.

1.6.4 Examples.

- Let V be the space of functions of $\mathbb{R}$ to itself. The map
$$\mathcal{A} : V \to \mathbb{R}, \quad f \mapsto f(0)$$
is linear is (because $\mathcal{A}(\lambda f + \mu g) = (\lambda f + \mu g)(0) = \lambda\, f(0) + \mu\, g(0) = \lambda \mathcal{A} f + \mu \mathcal{A} g$), so a linear function on V.

- Let W be the space of infinitely often differentiable functions on $\mathbb{R}$. The map
$$\mathcal{B} : W \to \mathbb{R}, \quad f \mapsto f'(0)$$
is also linear, so a linear function op W.

1.6.5 Example. If V is a real inner product space and $\underline{a} \in V$, then the map $\mathcal{A} : V \to \mathbb{R}$, defined by $\mathcal{A}\underline{x} = (\underline{a}, \underline{x})$ is a linear function. This follows from the linearity in the second argument of the inner product: $(\underline{a}, \lambda\, \underline{x} + \mu\, \underline{y}) = \lambda\, (\underline{a}, \underline{x}) + \mu\, (\underline{a}, \underline{y})$. The null space of this map is the orthoplement $< \underline{a} >^{\perp}$. The range is $\mathbb{R}$ if $\underline{a} \neq \underline{0}$ and is $\{\underline{0}\}$ if $\underline{a} = \underline{0}$.

1.6.6 The basis constructed in the following theorem is useful for the description of coordinates. In this respect this basis plays a role comparable with that of an orthonormal basis in an inner product space.

1.6.7 Theorem. *Let* $\{\underline{a}_1, \ldots, \underline{a}_n\}$ *be a basis of* V.

1. *There exists a unique basis* $\{\phi_1, \ldots, \phi_n\}$ *of* V^* *satisfying*
$$\begin{aligned} \phi_i(\underline{a}_j) &= 1 \quad \text{if } i = j, \\ &= 0 \quad \text{if } i \neq j. \end{aligned}$$
In particular one has $\dim(V) = \dim(V^*)$ *if* V *is finite dimensional.*

2. *If* $\underline{a} \in V$, *then* $(\phi_1(\underline{a}), \ldots, \phi_n(\underline{a}))$ *is the coordinate vector of* $\underline{a}$.

Proof.

1. According to Theorem 1.1.30 there exists for every $i = 1, \ldots, n$ a unique linear function $\phi_i : V \to \mathbb{R}$ with property mentioned in the theorem.

   It remains to prove that the system $\phi_1, \ldots, \phi_n$ is a basis of V^*. We first show that the system is independent. Suppose
$$\lambda_1 \phi_1 + \cdots + \lambda_n \phi_n = 0 \quad (= \text{the zero function}).$$

Now apply both sides to vector $\underline{a}_i$. On the left we find λ_i ; On the right we find 0. So $\lambda_i = 0$ for $i = 1, \ldots, n$ and the system is independent.

Next we show that every vector $\phi \in V^*$ belongs to the span $< \phi_1, \ldots, \phi_n >$. It is easy to check that $\phi(\underline{a}_1)\phi_1 + \phi(\underline{a}_2)\phi_2 + \cdots + \phi(\underline{a}_n)\phi_n$ and ϕ both give the same values for $\underline{a}_1, \ldots, \underline{a}_n$. But this means that $\phi = \phi(\underline{a}_1)\phi_1 + \phi(\underline{a}_2)\phi_2 + \cdots + \phi(\underline{a}_n)\phi_n$, again because of Theorem 1.1.30.

2. If $\underline{a} = \lambda_1 \underline{a}_1 + \cdots + \lambda_n \underline{a}_n$, then

$$\phi_i(\underline{a}) = \lambda_1 \phi_i(\underline{a}_1) + \cdots + \lambda_n \phi_i(\underline{a}_n) = \lambda_i$$

because of the properties of ϕ_i.

□

1.6.8 Definition. (Dual basis) The basis $\{\phi_1, \ldots, \phi_n\}$ of V^* is called the *dual basis* of $\{\underline{a}_1, \ldots, \underline{a}_n\}$.

1.6.9 Example. The dual basis of the standard basis $\{\underline{e}_1, \ldots, \underline{e}_n\}$ is denoted by $\{\underline{e}_1^*, \ldots, \underline{e}_n^*\}$. The dual basis of the basis $\alpha : \{\underline{e}_1 + \underline{e}_2,\ \underline{e}_1\}$ of $\mathbb{R}^2$ is $\alpha^* : \{\underline{e}_2^*,\ \underline{e}_1^* - \underline{e}_2^*\}$. It is now easy to compute the α–coordinates of the vector $(3, 7) = 3\,\underline{e}_1 + 7\,\underline{e}_2$: the first coordinate is $\underline{e}_2^*(3\,\underline{e}_1 + 7\,\underline{e}_2) = 7$, the second coordinate is $(\underline{e}_1^* - \underline{e}_2^*)(3\,\underline{e}_1 + 7\,\underline{e}_2) = 3 - 7 = -4$.

1.6.10 (Dual basis and the inverse matrix) Using our matrix techniques it is now easy to determine the dual basis of a basis of $\mathbb{K}^n$. As before $\{\underline{e}_1^*, \ldots, \underline{e}_n^*\}$ is the dual basis of the standard basis. Let $\{\underline{a}_1, \ldots, \underline{a}_n\}$ be any basis and collect these vectors als columns in a matrix $A = (a_{ij})$ (so this is the matrix ${}_\epsilon S_\alpha$). Every vector from the associated dual basis $\{\underline{a}_1^*, \ldots, \underline{a}_n^*\}$ can be written as a linear combination of $\underline{e}_1^*, \ldots, \underline{e}_n^*$, say, $\underline{a}_i^* = b_{i1}\,\underline{e}_1^* + \cdots + b_{in}\,\underline{e}_n^*$. We collect the b_{ij}'s a matrix B. It follows that

$$\underline{a}_i^*(\underline{a}_j) = b_{i1}a_{1j} + \cdots + b_{in}a_{nj}.$$

On the right hand side we see the (i, j)-th element of the product matrix BA . The left hand side is 1 if $i = j$ and 0 otherwise. This means that $I = BA$. So B is simply the inverse of A. The *rows* of B give the coordinates of the vectors from the dual basis.

1.6.11 Example. To find the dual basis of $\{\underline{e}_1 + \underline{e}_2,\ \underline{e}_1\}$ we invert the matrix

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

and find

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

So the dual basis is $\{\underline{e}_2^*$ (first row), $\underline{e}_1^* - \underline{e}_2^*$ (second row)$\}$.

1.6.12 (Determinant function) We can also use the dual basis to write down a determinant function. Let $\{\phi_1, \dots, \phi_n\}$ be the dual basis of the standard basis in $\mathbb{K}^n$ and define

$$D(\underline{a}_1, \dots, \underline{a}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, \phi_{\sigma(1)}(\underline{a}_1) \cdots \phi_{\sigma(n)}(\underline{a}_n)$$

(Here S_n denotes the set of $n!$ permutations of $\{1, 2, \dots, n\}$). We can check this using the properties of permutations, but we will not do that here. For $n = 2$ the expression simplifies to

$$\phi_1(\underline{a}_1)\phi_2(\underline{a}_2) - \phi_2(\underline{a}_1)\phi_1(\underline{a}_2) = a_{11}a_{22} - a_{21}a_{12}$$

if $\underline{a}_1 = (a_{11}, a_{21})$ and $\underline{a}_2 = (a_{12}, a_{22})$.

1.6.13 (Dual map) For a linear map $\mathcal{A} : V \to W$, we can define another linear map $\mathcal{A}^* : W^* \to V^*$ as follows:

$$\mathcal{A}^*(f) = f \circ \mathcal{A},$$

for every linear function $f : W \to \mathbb{K}$ from W^*. The map $\mathcal{A}^*$ is called *the dual map induced by* $\mathcal{A}$.

$$\begin{tikzcd}
V \arrow[rr, "\mathcal{A}"] \arrow[dr, "\mathcal{A}^* f"'] & & W \arrow[dl, "f"] \\
& \mathbb{K} &
\end{tikzcd}$$

When we represent $\mathcal{A}$ with a matrix $A = {}_{\beta}A_{\alpha}$ w.r.t. the bases $\alpha = \{\underline{a}_1, \dots, \underline{a}_n\}$ of V and $\beta = \{\underline{b}_1, \dots, \underline{b}_m\}$ of W, then it turns out that the matrix of $\mathcal{A}^*$ (w.r.t. the dual basis $\beta^* = \{\underline{b}_1^*, \dots, \underline{b}_m^*\}$ of W^* and $\alpha^* = \{\underline{a}_1^*, \dots, \underline{a}_n^*\}$ of V^*) is equal to $A^\top$. To see this we determine (i, j)–th element of the matrix A^* of $\mathcal{A}^*$. This requires finding the i–th α^*–coordinate of the image $\mathcal{A}^*(\underline{b}_j^*)$ of the j–th β^*–basis vector. If (assuming $\dim(V) = n$)

$$\mathcal{A}^*(\underline{b}_j^*) = \underline{b}_j^* \circ \mathcal{A} = \lambda_1 \underline{a}_1^* + \cdots + \lambda_n \underline{a}_n^*,$$

then we find the i–th coordinate by applying left and right hand side to the vector $\underline{a}_i$:

$$\underline{b}_j^* \circ \mathcal{A}(\underline{a}_i) = \underline{b}_j^*(\mathcal{A}(\underline{a}_i)).$$

Applying $\underline{b}_j^*$ to a vector gives the j–th coordinate, in this case the element A_{ji} of the matrix of $\mathcal{A}$. So it appears that $A^* = A^\top$.

1.7 Notes

The central notion when studying the connections between vector spaces is that of a linear map; these maps "respect" the linear structure in the sense that linear combinations of vectors are mapped to (the same) linear combinations of the image vectors, and linear subspaces are mapped to linear subspaces etc. They are the most obvious maps to study in the context of vector spaces and fortunately many relevante and important maps belong to this categorie (a.o. refelections, rotations, projections) to justify separate study. This of course does not mean that other, more general types of maps are not also relevant, but the other types are certainly more difficult to analyse. Quadratic maps for instance are extremely relevant and occur here only in the discussion about quadratic forms (Chapter 2) and inner products . In general it is very useful to study for a given structure, such as in our case vector space, those maps that "respect" or "keep" this structure. This theme is more systematically elaborated in the algebra courses. *Algebra*

To understand linear maps, the notions of eigenvalue and eigenvector turned out to be extremely useful. The notion eigenvalue has its origin in the context of systems of linear differential equations, a topic that will be treated in the last chapter.

In quantum mechanics linear maps play a mayor role, because physical quantities, like velocity and momentum are described by linear maps in vector spaces of functions. Eigenvalues and eigenvectors have a special physical interpretation. *Quantum–mechanics*

Another important role of linear maps should also be mentioned. Linear maps appear as first–order approximations of differentiable functions of several variables. If $f : \mathbb{R}^n \to \mathbb{R}^n$ is such a function and $p \in \mathbb{R}^n$, then the matrix of this first orde approximation in p w.r.t. the standard basis equals the $n \times n$ matrix (i, j)-th element $\dfrac{\partial f_i}{\partial x_j}$. This is also useful in the study of curved spaces. *Analysis 2* *Tensor–calculus and Differential–geometry*

The reason linear maps are well-understood is because (at least in the finite dimensional case) they can be described in termen of a single matrix, and because of the theory of eigenvalues and eigenvectors. In fact, all linear maps can be put in a "normal form" (Chapter 3), see also [5].

1.8 Exercises

§1

1 Check whether the following maps are linear. Determine, in the cases that the map is linear, the null space and the range and verify the dimension theorem.

a. $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\mathcal{A}(x_1, x_2) = (x_1 + x_2, x_1)$,

b. $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $\mathcal{A}(x_1, x_2) = x_1 x_2$,

c. $\mathcal{A}\colon \mathbb{C}^3 \to \mathbb{C}^2$ defined by $\mathcal{A}(x_1, x_2, x_3) = (x_1 + ix_2, 0)$,

d. $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by $\mathcal{A}(x_1, x_2, x_3) = (|x_3|, 0)$,

e. $\mathcal{I}\colon V \to V$ defined by $\mathcal{I}(\underline{x}) = \underline{x}$ ('the identity map'); here V is a vector space,

f. $\mathcal{A}\colon V \to V$ defined by $\mathcal{A}\underline{x} = \underline{x} + \underline{a}$; here V is a vector space and $\underline{a} \in V$ a fixed vector,

g. $\mathcal{A}\colon V \to \mathbb{R}$ defined by $\mathcal{A}\underline{x} = (\underline{x}, \underline{a})$; here V is a real inner product space and $\underline{a}$ a fixed vector.

2 In this exercise V is the vector space of functions from $\mathbb{R}$ to $\mathbb{R}$. Check whether the following maps are linear. Determine, in those cases that the map is linear, the null space and the range.

a. $\mathcal{A}\colon V \to \mathbb{R}$ defined by $\mathcal{A}f = f(1)$,

b. $\mathcal{A}\colon V \to \mathbb{R}$ defined by $\mathcal{A}f = (f(0))^2$,

c. $\mathcal{A}\colon V \to V$ defined by $(\mathcal{A}f)(x) = 1 + f(x)$ for all $x \in \mathbb{R}$,

d. $\mathcal{A}\colon V \to V$ defined by $(\mathcal{A}f)(x) = f(x^2)$ for all $x \in \mathbb{R}$,

e. $\mathcal{A}\colon V \to V$ defined by $(\mathcal{A}f)(x) = (x+1)f(x)$ for all $x \in \mathbb{R}$.

3 Let V be the vector space of infinitely often differentiable functions from $\mathbb{R}$ to $\mathbb{R}$. We define the map $\mathcal{A}\colon V \to V$ by the rule: $\mathcal{A}f$ is the function satisfying $(\mathcal{A}f)(x) = f'(x) + 3f(x)$ for all $x \in \mathbb{R}$.

a. Show that $\mathcal{A}$ is linear.

b. Determine the dimension and a basis of the null space of $\mathcal{A}$.

c. Determine an $f \in V$ for which $(\mathcal{A}f)(x) = \sin x$ for all $x \in \mathbb{R}$.

d. Determine the range of $\mathcal{A}$.

4 $\mathcal{P}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the orthogonal projection on the line $< (1, 2, 3) >$.

a. Determine an expression for $\mathcal{P}\underline{x}$ and determine $\mathcal{P}(1, 1, 2)$.

b. Determine the range and the null space of $\mathcal{P}$. Is the map $\mathcal{P}$ bijective, in other words, does $\mathcal{P}$ have an inverse?

c. Determine all solutions of the equation $\mathcal{P}\underline{x} = (2, 4, 6)$.

d. Determine all solutions of the equation $\mathcal{P}\underline{x} = (1, 2, 2)$.

e. What is the geometric description of the map $\mathcal{I} - \mathcal{P}\colon \mathbb{R}^3 \to \mathbb{R}^3$ (so $(\mathcal{I} - \mathcal{P})\underline{x} = \underline{x} - \mathcal{P}\underline{x}$)?

5 We are given a vector $\underline{a} \in \mathbb{R}^3$ with $|\underline{a}| = 1$. For every real number λ the map $\mathcal{A}_\lambda\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$\mathcal{A}_\lambda \underline{x} = \underline{x} + \lambda(\underline{x}, \underline{a})\underline{a}.$$

a. Prove that for every λ the map $\mathcal{A}_\lambda$ is linear.

b. For which λ is it true that $\mathcal{A}_\lambda^2 = \mathcal{A}_\lambda$? What is the geometric interpretation of $\mathcal{A}_\lambda$ for these λ?.

c. For which λ is it true that $\mathcal{A}_\lambda^2 = \mathcal{I}$ (the identity map)? What is a geometric interpretation of $\mathcal{A}_\lambda$.

d. For which λ does the null space of $\mathcal{A}_\lambda$ equal $\{\underline{0}\}$?

6 Which of the following statements is true respectively false. Motivate your answer.

a. There exists a linear map $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ with $\mathcal{A}(2, 3) = (1, 0)$.

b. There exists a linear map $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ with $\mathcal{A}(2, 3) = (1, 0)$ and $\mathcal{A}(4, 6) = (0, 1)$.

c. There exists a linear map $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ with $\mathcal{A}(2, 3) = (1, 0)$ and $\mathcal{A}(3, 2) = (0, 1)$.

d. If the range of the linear map $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ contains the vectors $(2,3)$ and $(3,2)$, then $\mathcal{A}$ has an inverse.

§2

7 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$\mathcal{A}(1,2,3) = \mathcal{A}(2,3,1) = \mathcal{A}(2,1,3) = (6,-36,30).$$

Determine:

a. the matrix of $\mathcal{A}$,

b. the dimension and a basis of the range of $\mathcal{A}$,

c. the dimension and a basis of the null space of $\mathcal{A}$,

d. all solutions of the equation $\mathcal{A}\underline{x} = (1,-6,5)$,

e. $\mathcal{A}(l)$ with l given by $l\colon \underline{x} = (0,-1,1) + \lambda(1,1,-2)$,

f. $\mathcal{A}(V)$ with V the plane given by $x + y + z = 1$.

8 The linear map $\mathcal{A}\colon \mathbb{R}^4 \to \mathbb{R}^3$ is given by

$$\begin{aligned} \mathcal{A}(1,0,0,0) &= (3,-1,1), \quad & \mathcal{A}(0,1,0,0) &= (1,3,1), \\ \mathcal{A}(0,0,1,0) &= (1,1,1), \quad & \mathcal{A}(0,0,0,1) &= (1,1,1). \end{aligned}$$

Determine:

a. the dimension and a basis of the range of $\mathcal{A}$,

b. the dimension and a basis of the null space of $\mathcal{A}$,

c. all solutions of the equation $\mathcal{A}\underline{x} = (1,1,1)$,

d. $\mathcal{A}(l)$ with l the line given by $\underline{x} = (1,1,0,0) + \lambda(2,1,0,0)$,

e. $\mathcal{A}^{-1}(V)$ with V the plane given by the equation $z = 1$.

9 The linear map $\mathcal{A}\colon \mathbb{C}^3 \to \mathbb{C}^3$ is given by the matrix

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

a. Compute $\mathcal{A}(i,1,1)$, $\mathcal{A}(0,1,1)$, $\mathcal{A}(1,0,-i)$, $\mathcal{A}(1,1,1-i)$.

b. Determine the null space and the range of $\mathcal{A}$.

c. Determine the image under the map $\mathcal{A}$ of the line in $\mathbb{C}^3$ with parameter representation $\underline{x} = (0,1,1) + \lambda(i,1,1)$.

d. Determine all solutions of the equation $\mathcal{A}\underline{x} = (0,1,1)$.

10 Determine in each of the following cases the matrix of the given map.

a. The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$, given by

$$\mathcal{A}(2,1,1) = (7,6,7), \quad \mathcal{A}(1,0,2) = (1,15,10), \quad \mathcal{A}(-1,2,2) = (-1,7,4).$$

b. The linear map $\mathcal{A}\colon \mathbb{R}^4 \to \mathbb{R}^3$, given by

$$\begin{gathered}\mathcal{A}(0,0,1,-1) = (0,0,0), \quad \mathcal{A}(1,1,0,-2) = (0,0,0),\\ \mathcal{A}(1,1,1,1) = (4,4,4), \quad \mathcal{A}(1,2,3,4) = (4,12,10).\end{gathered}$$

c. The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^2$, whose null space is the line $<(1,-3,-2)>$ and for which moreover $\mathcal{A}(1,0,0) = (3,2)$, $\mathcal{A}(1,1,1) = (4,3)$.

11 The linear map $\mathcal{P}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the projection on the plane $2x + y + z = 0$. The linear map $\mathcal{Q}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the projection on the line $<(2,1,1)>$.

a. Determine the matrix of $\mathcal{P}$.

b. Determine $\mathcal{P}(1,1,1)$ and $\mathcal{P}(1,-1,-1)$.

c. Determine the matrix of $\mathcal{Q}$.

12 The linear map $\mathcal{S}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the reflection in the plane $2x - y + 3z = 0$.

a. Determine $\mathcal{S}(2,-1,3)$, $\mathcal{S}(1,2,0)$, $\mathcal{S}(0,3,1)$.

b. Determine the matrix of $\mathcal{S}$.

13 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

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

a. Determine the null space of $\mathcal{A}$. Prove that $\mathcal{A}$ has and inverse.

b. Determine the matrix of the inverse map $\mathcal{A}^{-1}$.

§3

14 In $\mathbb{R}^3$ we consider two bases α and β:

$$\alpha = \{(1,2,3),(2,3,1),(2,1,3)\}, \quad \beta = \{(3,5,4),(4,4,4),(5,6,7)\}.$$

As usual $\varepsilon = \{(1,0,0),(0,1,0),(0,0,1)\}$ is the standard basis.

a. Show that both α and β are indeed bases of $\mathbb{R}^3$.

b. Determine the transition matrices ${}_{\alpha}S_{\varepsilon}$ and ${}_{\beta}S_{\varepsilon}$.

c. Determine the coordinates $\alpha(3,3,6)$ and $\beta(3,3,6)$.

d. Determine the transition matrix ${}_{\beta}S_{\alpha}$ from α to β. Use this to verify the relation $\beta(3,3,6) = {}_{\beta}S_{\alpha}\alpha(3,3,6)$.

15 In the space V of real polynomials of degree at most 2 we are given the two bases $\alpha = \{1, x, x^2\}$ and $\beta = \{x^2+x, x+1, x^2+1\}$.

a. Show that β is indeed a basis of V.

b. Determine the transition matrix ${}_{\beta}S_{\alpha}$ from α to β.

c. Determine the coordinate vector $\beta(p)$ of the polynomial $p = 4x^2 - 6x + 2$.

16 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

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

Determine the matrix A_{α} of $\mathcal{A}$ w.r.t. the basis

$$\alpha = \{(2,1,1),(1,0,2),(-1,2,3)\}.$$

17 a. $\mathcal{P}\colon \mathbb{R}^2 \to \mathbb{R}^2$ is the orthogonal projection on the line $x - y = 0$. We are given the bases $\varepsilon = \{(1,0),(0,1)\}$ (the standard basis) and $\alpha = \{(1,-1),(1,1)\}$. Determine the transition matrices ${}_{\varepsilon}S_{\alpha}$ and ${}_{\alpha}S_{\varepsilon}$. Determine the matrix of $\mathcal{P}$ w.r.t. the basis α.

b. $\mathcal{P}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the projection on the plane $x+y-z=0$. We are given the bases

$$\begin{aligned} \alpha &= \{(1,1,-1),(1,0,1),(1,-1,0)\} \quad \text{and} \\ \beta &= \{(2,1,0),(2,-1,1),(2,0,-1)\}. \end{aligned}$$

Determine the transition matrices ${}_\alpha S_\beta$ and ${}_\beta S_\alpha$. Determine the matrix of $\mathcal{P}$ w.r.t. each of the two bases.

18 Let V be the space of real polynomials of degree at most 2. Furthermore $\mathcal{A}\colon V \to V$ maps a polynomial $p(x)$ to the derivative of $(x+1)p(x)$: $\mathcal{A}p = (gp)'$ (here $g(x) = x+1$).

a. Show that the map $\mathcal{A}$ is linear.

b. Determine the matrix A_α of $\mathcal{A}$ w.r.t. the basis $\alpha = \{1, x, x^2\}$.

c. Determine the matrix A_β of $\mathcal{A}$ w.r.t. the basis $\beta = \{x^2+x, x+1, x^2+1\}$.

§4

19 Determine eigenvalues and eigenspaces of the linear maps $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by the following matrices. Find out whether $\mathbb{R}^2$ has a basis consisting of eigenvectors and, if yes, determine such a basis and give the matrix of $\mathcal{A}$ w.r.t. this basis.

a. $\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$,

b. $\begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix}$,

c. $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$,

d. $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$,

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

20 Determine the eigenvalues and the corresponding eigenspaces of the linear maps $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by the following matrices and find out whether $\mathbb{R}^3$ has a basis consisting of eigenvectors of $\mathcal{A}$ and, if yes, determine such a basis and give the matrix of $\mathcal{A}$ w.r.t. this basis.

a. $A = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix}$,

b. $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$,

c. $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$,

d. $A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ -2 & 2 & 1 \end{pmatrix}$,

e. $A = \begin{pmatrix} 1 & 6 & -4 \\ 6 & 2 & -2 \\ -4 & -2 & -3 \end{pmatrix}$,

f. $A = \dfrac{1}{2}\begin{pmatrix} 3 & 0 & -1 \\ 2 & 2 & -2 \\ 3 & 0 & -1 \end{pmatrix}$.

21 Determine the eigenvalues and the corresponding eigenspaces of the linear maps $\mathcal{A}\colon \mathbb{C}^3 \to \mathbb{C}^3$ given by the following matrices and find out whether $\mathbb{C}^3$ has a basis consisting of eigenvectors of $\mathcal{A}$.

a. $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{pmatrix}$,

b. $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1+i & 1 \\ 0 & -i & 0 \end{pmatrix}$,

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

22 Let $h(x) = x^2 + x$ in the space V of real polynomials of degree at most 2. The map $\mathcal{A}\colon V \to V$ is given by $\mathcal{A}p = (hp)''$ (second derivative).

a. Show that the map $\mathcal{A}$ is linear.

b. Determine the eigenvalues and the corresponding eigenspaces of $\mathcal{A}$.

c. What is the form of the matrix of $\mathcal{A}$ w.r.t. a basis of eigenvectors?

23 W.r.t. the basis $\alpha = \{\underline{a}_1, \underline{a}_2, \underline{a}_3\}$ the linear map $\mathcal{A}\colon \mathbb{C}^3 \to \mathbb{C}^3$ is given by the matrix

$$A_\alpha = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -i \\ 1 & -i & 0 \end{pmatrix}.$$

a. Determine the eigenvalues and the corresponding eigenspaces of $\mathcal{A}$.

b. Determine the null space and the range of $\mathcal{A}$.

c. Show that there is a basis of $\mathbb{C}^3$ consisiting of eigenvectors of $\mathcal{A}$ and determine the matrix of $\mathcal{A}$ w.r.t. this basis.

24 In this exercise V is a real or complex vector space. Show that:

a. If the linear map $\mathcal{A}\colon V \to V$ satisfies $\mathcal{A}^2 = \mathcal{A}$ and if λ is an eigenvalue of $\mathcal{A}$, then $\lambda = 0$ or $\lambda = 1$. Give a (non trivial) example.

b. If the linear map $\mathcal{A}\colon V \to V$ satisfies $\mathcal{A}^2 = \mathcal{I}$ and if λ is an eigenvalue of $\mathcal{A}$, then $\lambda = 1$ or $\lambda = -1$. Give again a (non trivial) example.

c. If the linear map $\mathcal{A}\colon V \to V$ has an inverse and λ is an eigenvalue of $\mathcal{A}$, then:

i) $\lambda \neq 0$,
ii) $1/\lambda$ is an eigenvalue of $\mathcal{A}^{-1}$,
iii) $E_{1/\lambda}(\mathcal{A}^{-1}) = E_\lambda(\mathcal{A})$.

25 The linear map $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 8 & 3 \end{pmatrix}.$$

a. The map $\mathcal{A}$ has two eigenvalues, say λ_1, λ_2. Determine these and determine the corresponding eigenspaces $E_{\lambda_1},\ E_{\lambda_2}$ of $\mathcal{A}$.

b. Let α be a basis of $\mathbb{R}^2$ consisting of eigenvectors of $\mathcal{A}$. Determine the matrix A_α of $\mathcal{A}$ w.r.t. the basis α.

c. What is the connection between A^k and $(A_\alpha)^k$ (here k is a positive integer)?

d. We construct a sequence of vectors $\underline{x}_0, \underline{x}_1, \underline{x}_2, \ldots$ in $\mathbb{R}^2$ as follows:

$$\underline{x}_0 = (2,2), \quad \underline{x}_{k+1} = \mathcal{A}\underline{x}_k \quad (k = 0,1,2,\ldots).$$

Determine the coordinate vectors $\alpha(\underline{x}_k)$ $(k = 0,1,2,\ldots)$.

e. Determine $\underline{x}_k$ for $k = 0,1,2,\ldots$.

§5

26 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

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

Which of the following linear subspaces of $\mathbb{R}^3$ are invariant under $\mathcal{A}$? Motivate your answer.

a. $<(1,-1,1)>$,
b. $<(1,-1,0),(1,1,-2)>$,
c. $x - y - z = 0$,
d. $<(1,1,1),(0,0,1)>$.

27 In the vector space of differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ we consider the linear subspace $U = <\cos x, \sin x, x, 1>$. The linear map $\mathcal{D}\colon U \to U$ is given by differentiation:

$$(\mathcal{D}f)(x) = f'(x).$$

a. Which of the following subspaces of U are invariant under $\mathcal{D}$?

(i). $< 1 >$, (iii). $< \cos x >$,
(ii). $< 1, x >$, (iv). $< \cos x, \sin x >$.

b. Determine all 1-dimensional linear subspaces of U invariant under $\mathcal{D}$.

28 In the 3-dimensional real inner product space V we have the vectors $\underline{a}$ and $\underline{b}$ satisfying $||\underline{a}|| = ||\underline{b}|| = 1$, $(\underline{a}, \underline{b}) = 1/2$. The linear map $\mathcal{A}\colon V \to V$ is defined by $\mathcal{A}\underline{x} = (\underline{x}, \underline{b})\underline{a} - (\underline{x}, \underline{a})\underline{b}$.

a. Show that $\underline{a}, \underline{b}$ is an independent system.

b. Show that $U =< \underline{a}, \underline{b} >$ and $U^{\perp}$ are invariant under $\mathcal{A}$.

c. Determine the matrix of $\mathcal{A}$ w.r.t. of a basis containing $\underline{a}$ and $\underline{b}$.

d. Prove that $\mathcal{A}$ is not diagonalisable.

29 The linear map $\mathcal{A}\colon \mathbb{R}^5 \to \mathbb{R}^5$ maps $(1, 2, 1, 1, 2)$ to $(2, 4, 3, 3, 3)$ and $(1, 2, 2, 2, 1)$ to $(5, 10, 6, 6, 9)$.

a. Show that $W =< (1, 2, 1, 1, 2), (1, 2, 2, 2, 1) >$ is invariant under $\mathcal{A}$.

b. Determine the eigenvalues and the corresponding eigenspaces of the restriction $\mathcal{A}\colon W \to W$.

c. Show that there is a basis of W consisting of eigenvectors of $\mathcal{A}\colon W \to W$ and determine the matrix of $\mathcal{A}\colon W \to W$ w.r.t. this basis.

30 V is a vector space and $\mathcal{A}\colon V \to V$ a linear map. Prove the following statements:

a. If U is a subspace of V that is invariant is under $\mathcal{A}$, then so is $\mathcal{A}(U)$.

b. If λ is an eigenvalue of $\mathcal{A}$ with eigenspace E_λ, then E_λ is invariant under $\mathcal{A}$.

c. If every linear subspace of V is invariant under $\mathcal{A}$, then $\mathcal{A}$ has a basis consisting of eigenvectors. Much stronger: $\mathcal{A}$ is a multiple of the identity map $\mathcal{I}$.

31 For every real λ the linear map $\mathcal{A}_\lambda\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

$$A_\lambda = \begin{pmatrix} 11-\lambda & 6-\lambda & -4-\lambda \\ -4 & -4 & 16 \\ 3\lambda+5 & 10-2\lambda & -2\lambda \end{pmatrix}.$$

a. For which λ does $\mathcal{A}_\lambda$ have an inverse?

b. For which λ is the subspace $< (1, 2, 2) >$ invariant under $\mathcal{A}_\lambda$?

c. Same question for the plane $x + y + z = 0$.

32 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

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

a. Determine the eigenvectors and the corresponding eigenspaces of $\mathcal{A}$.

b. Determine, using a complex root of the characteristic equation a 2-dimensional subspace U of $\mathbb{R}^3$ invariant under $\mathcal{A}$.

c. Determine the matrix of $\mathcal{A}$ w.r.t. a basis that consists of an eigenvector of $\mathcal{A}$ completed with a basis of U.

§6

33 In $\mathbb{R}^2$ we work with the basis $\alpha : \{(2, 1),\ (3, 1)\}$.

a. Determine the dual basis of α.

b. Determine, using the dual basis, the coordinates w.r.t. α of the vector $(3, 4)$.

34 Let V be a real inner product space of dimension n. For every vector $\underline{a} \in V$ we have a linear function $\mathcal{A}_{\underline{a}} : V \to \mathbb{R}$, $\mathcal{A}_{\underline{a}}\, \underline{x} = (\underline{a}, \underline{x})$ (example 1.6.5).

a. Show that the map $\Phi : V \to V^*$, defined by $\Phi(\underline{a}) = \mathcal{A}_{\underline{a}}$ is linear.

b. Show that $\mathcal{N}(\Phi) = \{\underline{0}\}$ and conclude that the map Φ is an isomorphism (has an inverse).

c. Conclude that for every linear function $f : V \to \mathbb{R}$ there exists a unique vector $\underline{a} \in V$ such that $f(\underline{x}) = (\underline{a}, \underline{x})$ for all $\underline{x}$.

d. If $\underline{a}_1, \dots, \underline{a}_n$ is an orthonormal basis is of V, then $\Phi(\underline{a}_1), \dots, \Phi(\underline{a}_n)$ is the dual basis. Prove this.

35 Let W be a subspace of the vector space V and $\mathcal{A} : V \to V$ a linear map with range W satisfying $\mathcal{A}\underline{w} = \underline{w}$ for every $\underline{w} \in W$. Such a map is called a (not necessarily orthogonal) projection; the usual orthogonal projections from example 1.1.4 are examples of this.

a. Let $U := \mathcal{N}(\mathcal{A})$. Show that $U \cap W = \{\underline{0}\}$.

b. Let $\underline{x} \in V$. Show that $\underline{x} - \mathcal{A}\underline{x} \in U$. Conclude that every vector from V can be written as the sum of a vector from U and one from W.

c. If $\{\underline{u}_1, \ldots, \underline{u}_k\}$ is a basis of U and $\{\underline{w}_1, \ldots, \underline{w}_m\}$ a basis of W, then $\{\underline{u}_1, \ldots, \underline{u}_k, \underline{w}_1, \ldots, \underline{w}_m\}$ is a basis of V. Show this.

d. Let $\{\underline{u}_1^*, \ldots, \underline{u}_k^*, \underline{w}_1^*, \ldots, \underline{w}_m^*\}$ be the dual basis of the basis from the previous part. Prove that $\mathcal{A}\underline{x} = \underline{w}_1^*(\underline{x})\, \underline{w}_1 + \cdots + \underline{w}_m^*(\underline{x})\, \underline{w}_m$ (cf. the formula for an orthogonal projection in example 1.1.4).

1.8.1 Exercises on exam level

36 The matrix A of the linear map $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^4$ equals

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

a. Determine the null space of $\mathcal{A}$ and the dimension of the range of $\mathcal{A}$.

b. Describe the range of $\mathcal{A}$ using an equation of the form $a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = b$.

c. Determine a basis for the image under $\mathcal{A}$ of the plane

$$\{(y_1, y_2, y_3) \in \mathbb{R}^3 \mid y_1 + y_2 + y_3 = 0\}.$$

d. Determine the (complete) inverse image $\mathcal{A}^{-1}(m)$ of the line m in $\mathbb{R}^4$ with parameter representation

$$m : \underline{x} = (3, -1, -2, -2) + \lambda(1, 1, 0, -2).$$

37 The linear map $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^3$ is determined by:

$$\begin{aligned} \mathcal{A}(1,0,2) &= (4,1,-2), \\ \mathcal{A}(2,1,0) &= (7,-2,-1), \\ \mathcal{A}(0,2,1) &= (7,1,-3). \end{aligned}$$

a. Determine the range and the null space of $\mathcal{A}$. Does $\mathcal{A}$ have an inverse?

b. Determine the image of the plane $V : \underline{x} = \lambda(1, 1, 0) + \mu(0, -2, 1)$.

c. Determine the (complete) inverse image $\mathcal{A}^{-1}(m)$ of the line $m : \underline{x} = (0, 4, 3) + \rho(1, 2, 4)$.

38 In the real 3-dimensional vector space V with basis $\underline{a}$, $\underline{b}$, $\underline{c}$ the linear map $\mathcal{A} : V \to V$ is specified by

$$\begin{aligned} \mathcal{A}\underline{a} &= 3\underline{a} + 4\underline{b} + 3\underline{c}, \\ \mathcal{A}\underline{b} &= -\underline{a} - 3\underline{b} - 4\underline{c}, \\ \mathcal{A}\underline{c} &= \underline{a} + \underline{b} + 2\underline{c}. \end{aligned}$$

a. Show that the subspace $U =< \underline{a}+\underline{b}, \underline{a}-\underline{c} >$ is invariant under $\mathcal{A}$. What is the 2×2 matrix of the restriction of $\mathcal{A}$ to U w.r.t. the basis $\underline{a}+\underline{b}, \underline{a}-\underline{c}$ of U?

b. Determine the matrix of $\mathcal{A}$ w.r.t. the basis $\underline{a} + \underline{b}, \underline{a} - \underline{c}, \underline{c}$ of V.

c. Determine the eigenvalues and eigenspaces of the map $\mathcal{A}$.

Chapter 2

Orthogonal and symmetric maps

2.1 Orthogonal maps

In the previous chapter we have studied linear maps $\mathcal{A} : V \to V$ where V is a vector space. So far we have not used any special properties of this vector space. In this chapter V will be a real inner product space. So we have an inner product and therefore also the notions length and angle and in particular perpendicularity. We will study two classes of linear maps that have nice properties with respect to the inner product: orthogonal and symmetric maps. In both cases we first discuss the theoretical side, and then look at the role played by the matrices that describe these maps. In both cases it is possible to determine the possibilities in great detail. We will restrict ourselves to *finite dimensional, real* inner product spaces. In this section we discuss

- the notion orthogonal linear map;
- the notion orthogonal matrix, as well as the connection with orthogonal maps, and
- the classification of orthogonal maps, with emphasis on dimensions 2 and 3.

2.1.1 Definition. (Orthogonal map) Let V be a real inner product space. A linear map $\mathcal{A} : V \to V$ is called *orthogonal* if $\|\mathcal{A}\underline{x}\| = \|\underline{x}\|$ for all $\underline{x} \in V$. In other words: a linear map $\mathcal{A} : V \to V$ is orthogonal if the *length is invariant* under $\mathcal{A}$.

2.1.2 Example. For the identity map $\mathcal{I} : V \to V$ one has $\mathcal{I}\underline{x} = \underline{x}$ for every vector $\underline{x}$. So certainly $\|\mathcal{I}\underline{x}\| = \|\underline{x}\|$. So the identity map is orthogonal.

In the same way $-\mathcal{I}$ is orthogonal. For $\lambda \neq \pm 1$ however $\lambda\mathcal{I}$ is not orthogonal.

2.1.3 Example. The linear map $\mathcal{A} : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\mathcal{A}(x, y) = \frac{1}{5}(3x + 4y, -4x + 3y)$$

is orthogonal because

$$\begin{aligned} \|\mathcal{A}(x,y)\|^2 &= \tfrac{1}{25}((3x+4y)^2 + (-4x+3y)^2) \\ &= \tfrac{1}{25}(9x^2 + 16y^2 + 24xy + 16x^2 + 9y^2 - 24xy) \\ &= x^2 + y^2 = \|(x,y)\|^2. \end{aligned}$$

2.1.4 Example. (Orthogonal reflection) Let $\underline{a}$ be a vector of length 1 in the inner product space V. The linear map $\mathcal{A} : V \to V$ is given by the rule $\mathcal{A}\underline{x} = \underline{x} - 2(\underline{x}, \underline{a})\underline{a}$. This map is orthogonal. To show this we expand $\|\mathcal{A}\underline{x}\|^2$:

$$\begin{aligned} (\mathcal{A}\underline{x}, \mathcal{A}\underline{x}) &= (\underline{x} - 2(\underline{x},\underline{a})\underline{a}, \underline{x} - 2(\underline{x},\underline{a})\underline{a}) \quad = \\ &= (\underline{x},\underline{x}) - 4(\underline{x},\underline{a})(\underline{x},\underline{a}) + 4(\underline{x},\underline{a})^2(\underline{a},\underline{a}) \quad = \quad (\underline{x},\underline{x}) \ , \end{aligned}$$

from which the claim immediately follows. From a geometric point of view it is not surprising that this map is orthogonal, because it is a (perpendicular, or orthogonal) reflection in the subspace $< \underline{a} >^{\perp}$. Start moving from $\underline{x}$ in the direction of $\underline{a}$ until you reach $< \underline{a} >^{\perp}$: so intersect the line $\underline{x} + \lambda\underline{a}$ with $< \underline{a} >^{\perp}$. This corresponds to $\lambda = -1$ and the vector $\underline{x} - (\underline{x}, \underline{a})\,\underline{a}$. To complete the reflection we have to subtract the vector $(\underline{x}, \underline{a})\,\underline{a}$ from $\underline{x}$ twice.

2.1.5 Example. If $\mathcal{P}$ is the orthogonal projection on a line $< \underline{a} >$ then the map $\mathcal{P}$ is not orthogonal: choose a vector $\underline{b}$ perpendicular to $\underline{a}$ different from the zero vector. Then $\mathcal{P}\underline{b} = \underline{0}$ so $0 = \|\mathcal{P}\underline{b}\| < \|\underline{b}\|$. (This is of course only possible if $\dim V > 1$).

2.1.6 In a real inner product space we have

$$(\underline{x} + \underline{y}, \underline{x} + \underline{y}) = (\underline{x}, \underline{x}) + 2(\underline{x}, \underline{y}) + (\underline{y}, \underline{y}) \ ,$$

so that

$$(\underline{x}, \underline{y}) = \frac{1}{2}\left((\underline{x} + \underline{y}, \underline{x} + \underline{y}) - (\underline{x}, \underline{x}) - (\underline{y}, \underline{y})\right)$$

(the so called *polarisation formula*). Inner products between vectors can therefore be expressed as combinations of functions of the length of the vectors. As a consequence we may also define orthogonality of a map by the invariance of the inner product; another consequence is that the mutual inner products between vectors don't change if we replace them by their images. The next theorem characterises the notion orthogonal map in four equivalent ways. The use of this will become clear when we discuss matrices of orthogonal maps.

2.1.7 Theorem. *Let V be a finite dimensional real inner product space. For a linear map $\mathcal{A} : V \to V$ the following statements are equivalent:*

1. *$\mathcal{A}$ is orthogonal, that is, $\|\mathcal{A}\underline{x}\| = \|\underline{x}\|$ for all $\underline{x} \in V$, in other words, the length is invariant under $\mathcal{A}$.*
2. *$(\mathcal{A}\underline{x}, \mathcal{A}\underline{y}) = (\underline{x}, \underline{y})$ for all $\underline{x}, \underline{y} \in V$, in other words, the inner product is invariant under $\mathcal{A}$.*
3. *For every orthonormal system $\underline{a}_1, \ldots, \underline{a}_n$ in V also $\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_n$ is an orthonormal system, in other words orthonormal systems are mapped to orthonormal systems.*
4. *There exists an orthonormal basis $\{\underline{a}_1, \ldots, \underline{a}_m\}$ of V such that also $\{\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_m\}$ is an orthonormal basis of V.*

Proof. It suffices to establish the following chain of implications: 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 1).

1) $\Rightarrow$ 2): Statement 1 means that for all $\underline{x} \in V$ we have $(\mathcal{A}\underline{x}, \mathcal{A}\underline{x}) = (\underline{x}, \underline{x})$. Take $\underline{x}$ and $\underline{y}$ in V. Then we have

$$(\mathcal{A}(\underline{x}+\underline{y})\,,\ \mathcal{A}(\underline{x}+\underline{y})) = (\underline{x}+\underline{y}, \underline{x}+\underline{y}), \quad \text{and hence}$$

$$(\mathcal{A}\underline{x}, \mathcal{A}\underline{x}) + 2(\mathcal{A}\underline{x}, \mathcal{A}\underline{y}) + (\mathcal{A}\underline{y}, \mathcal{A}\underline{y}) = (\underline{x}, \underline{x}) + 2(\underline{x}, \underline{y}) + (\underline{y}, \underline{y}).$$

Since $(\mathcal{A}\underline{x}, \mathcal{A}\underline{x}) = (\underline{x}, \underline{x})$ and $(\mathcal{A}\underline{y}, \mathcal{A}\underline{y}) = (\underline{y}, \underline{y})$ it follows that $(\mathcal{A}\underline{x}, \mathcal{A}\underline{y}) = (\underline{x}, \underline{y})$. So if the length is invariant under $\mathcal{A}$, then also the inner product.

2) $\Rightarrow$ 3): Let $\underline{a}_1, \ldots, \underline{a}_n$ be an orthonormal system. Then the inner product $(\mathcal{A}\underline{a}_i, \mathcal{A}\underline{a}_j) = (\underline{a}_i, \underline{a}_j) = 1$ if $i = j$, and 0 if $i \neq j$. So also $\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_n$ is an orthonormal system.

3) $\Rightarrow$ 4): Suppose that $\dim V = m$. Choose an orthonormal basis $\{\underline{a}_1, \ldots, \underline{a}_m\}$ of V (such a basis exists for example by the construction of Gram–Schmidt). According to 3) the system $\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_m$ is again orthonormal and therefore independent. Since $m = \dim V$ this system is a basis of V.

4) $\Rightarrow$ 1): Let $\{\underline{a}_1, \ldots, \underline{a}_m\}$ be an orthonormal basis as meant in 4), so $\{\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_m\}$ is again an orthonormal basis. Let moreover $\underline{x} \in V$. If $\underline{x} = \lambda_1\underline{a}_1 + \cdots + \lambda_m\underline{a}_m$, then $\mathcal{A}\underline{x} = \lambda_1\mathcal{A}\underline{a}_1 + \cdots + \lambda_m\mathcal{A}\underline{a}_m$. Now apply Pythagoras' Theorem to both expressions (the square of the length of $\lambda_i\underline{a}_i$ and of $\lambda_i\mathcal{A}\underline{a}_i$ is λ_i^2):

$$\|\underline{x}\|^2 = \lambda_1^2 + \cdots + \lambda_m^2 \quad \text{and} \quad \|\mathcal{A}\underline{x}\|^2 = \lambda_1^2 + \cdots + \lambda_m^2,$$

from which it immediately follows that $\|\mathcal{A}\underline{x}\| = \|\underline{x}\|$. □

2.1.8 Theorem. *Let V be a finite dimensional real inner product space.*

1. *If $\mathcal{A} : V \to V$ and $\mathcal{B} : V \to V$ are orthogonal, then also $\mathcal{AB} : V \to V$ is orthogonal.*
2. *If $\mathcal{A} : V \to V$ is orthogonal, then $\mathcal{A}$ is invertible and also $\mathcal{A}^{-1}$ is orthogonal.*

Proof. 1) For all $\underline{x} \in V$ we have $\|\mathcal{AB}\underline{x}\| = \|\mathcal{B}\underline{x}\| = \|\underline{x}\|$ because $\mathcal{A}$ and $\mathcal{B}$ are orthogonal, so also $\mathcal{AB}$ is orthogonal.

2) First we show that $\mathcal{A}$ is invertible. Apply the definition of orthogonal map to a vector $\underline{x}$ from the null space of $\mathcal{A}$. It follows from $0 = \|\mathcal{A}\underline{x}\| = \|\underline{x}\|$ that $\underline{x} = \underline{0}$. From 1.1.36 it now follows that $\mathcal{A}$ has an inverse. For all $\underline{x} \in V$ we have $\underline{x} = \mathcal{A}\mathcal{A}^{-1}\underline{x}$. Since $\mathcal{A}$ is orthogonal this implies $\|\underline{x}\| = \|\mathcal{A}\mathcal{A}^{-1}\underline{x}\| = \|\mathcal{A}^{-1}\underline{x}\|$ so $\mathcal{A}^{-1}$ is also orthogonal. □

2.1.9 In infinite dimensional inner product spaces there are orthogonal maps that are not invertible.

2.1.10 From now on we will concentrate on orthogonal maps in $\mathbb{R}^n$ and their matrices. This leads to the notion of an *orthogonal matrix*. Because the standard basis is an orthonormal basis of $\mathbb{R}^n$ (with the standard inner product) we get from Theorem 2.1.7:

2.1.11 Theorem. *A linear map $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if and only if $\mathcal{A}\underline{e}_1, \ldots, \mathcal{A}\underline{e}_n$ is an orthonormal system.*

2.1.12 To check whether such a map $\mathcal{A}$ is orthogonal, we may verify whether the columns of the matrix A of $\mathcal{A}$ form an orthonormal system. One way to do this, is to use matrix multiplication: element ij of the product $A^\top A$ equals the inner product of the i–th and the j–th column of A. Checking whether the columns of A form an orthonormal system, then amounts to verifying that $A^\top A = I$. If the map is orthogonal, then $A^\top$ is the inverse of A and the matrix of $\mathcal{A}^{-1}$. Since this inverse map is also orthogonal is, we see that the columns of $A^\top$ and therefore the rows of A form an orthonormal system. Time for a definition and a theorem concerning these observations.

2.1.13 Definition. (Orthogonal matrix) A real $n \times n$ matrix A is called *orthogonal* if the columns form an orthonormal system in $\mathbb{R}^n$.

Summarizing:

2.1.14 Theorem. *Let $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map with matrix A. The following are equivalent:*

- *$\mathcal{A}$ is orthogonal.*
- *The matrix A is an orthogonal matrix.*
- *$A^\top A = I$.*
- *The columns of A form an orthonormal system.*
- *The rows of A form an orthonormal system.*

2.1.15 Example. Every column of the matrix

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

has length 1 and every two different columns have inner product 0; the matrix A is therefore orthogonal. (Also the rows have length 1 and mutual inner product 0.) To determine the inverse of A we only have to transpose:

$$A^{-1} = A^\top = \frac{1}{3}\begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ -2 & 2 & 1 \end{pmatrix}.$$

2.1.16 We will now discuss the role of matrices for general orthogonal maps. Let V be a real inner product space and α and β two orthonormal bases for V. It follows that the length of a vector and its coordinate vector are equal. This means in particular that the length of the α– and the β–coordinate vector of a vector from V are equal. The change of coordinates map $\beta\alpha^{-1}$ therefore leaves lengths invariant and is therefore an orthogonal map; and hence the associated transition matrix ${}_\beta S_\alpha$ is also orthogonal.

2.1.17 Theorem. *If α and β are two orthonormal bases in a real inner product space, then the transition matrix ${}_\beta S_\alpha$ is orthogonal and we have*

$$ {}_\alpha S_\beta = {}_\beta S_\alpha^\top . $$

It is important that α and β are *orthonormal* bases, the theorem does not hold without this condition!

2.1.18 Let α be an orthonormal basis for an n-dimensional real inner product space V and consider a linear map $\mathcal{A} : V \to V$. The matrix A_α of $\mathcal{A}$ w.r.t. this basis is the matrix of the linear map $\alpha\mathcal{A}\alpha^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. If $\mathcal{A}$ if orthogonal, then also $\alpha\mathcal{A}\alpha^{-1}$ (since all maps in this product leave lengths invariant), so the matrix A_α is orthogonal. Conversely, if the matrix A_α is orthogonal, then the map $\alpha\mathcal{A}\alpha^{-1}$ is orthogonal and therefore also $\mathcal{A} = \alpha^{-1}(\alpha\mathcal{A}\alpha^{-1})\alpha$. This gives us the following variation of 2.1.14:

2.1.19 Theorem. *Let α be an orthonormal basis for a real inner product space V. The linear map $\mathcal{A} : V \to V$ is orthogonal if and only if the matrix A_α is orthogonal.*

2.1.20 Example. Back to example 2.1.4. To determine whether the map $\mathcal{A}$ is orthogonal, we can also determine the matrix w.r.t. a cleverly chosen orthonormal basis. Choose an orthonormal basis $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ whose first vector is $\underline{a}$. Then $\mathcal{A}\underline{a}_1 = -\underline{a}_1$, and $\mathcal{A}\underline{a}_i = \underline{a}_i$ for $i = 2, \ldots, n$. The matrix is therefore a diagonal matrix with on the diagonal $-1, 1, \ldots, 1$. It is not difficult to check that this matrix is orthogonal.

2.1.21 (Classification of orthogonal maps) In the rest of this section we analyse the possibilities for orthogonal maps in the different dimensions, in particular in dimensions 1, 2 and 3. It turns out that orthogonal maps consist of rotations, reflections and combinations of them. W.r.t. an appropriate basis we get a matrix representation that is easy to interpret. We will now go into this in more detail.

2.1.22 If A is an orthogonal matrix, then $A^\top A = I$, so $\det(A^\top)\det(A) = \det(I) = 1$. Since $\det(A^\top) = \det(A)$ if follows that $\det(A)^2 = 1$ and as a consequence $\det(A) = \pm 1$. W.r.t. an arbitrary orthonormal basis α an orthogonal map $\mathcal{A} : V \to V$ has an orthogonal matrix A. So for this matrix we have $\det(A_\alpha) = 1$ or $\det(A_\alpha) = -1$. The value of this determinant is, according to Theorem 1.4.16 independent of the basis α. So orthogonal maps come in two types:

- The *directly orthogonal* maps, with $\det(A_\alpha) = 1$ w.r.t. any orthonormal basis α,
- the *indirectly orthogonal* maps, with $\det(A_\alpha) = -1$ w.r.t. any orthonormal basis α.

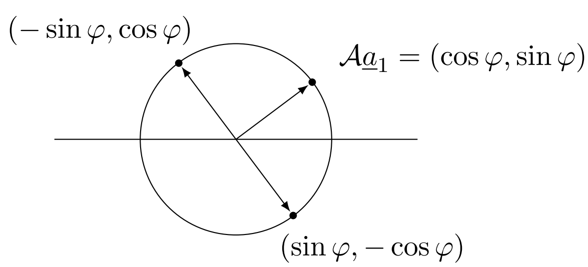
2.1.23 The characteristic polynomial $\det(A_\alpha - \lambda I)$ has n (complex) roots $\lambda_1, \ldots, \lambda_n$. The real roots are eigenvalues and since an orthogonal map leaves length invariant, every real root is ± 1. A non-real root a comes together with its complex conjugate $\bar{a}$, and $a\,\bar{a} = |a|^2 > 0$. In Theorem 1.4.18 we have already seen that

$$\det(A_\alpha) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

It follows that if $\det(A_\alpha) = -1$, so if $\mathcal{A}$ is indirectly orthogonal, then there must be an eigenvalue -1. We shall see soon that this means that $\mathcal{A}$ "contains" a reflection in $E_{-1}^{\perp}$.

2.1.24 (Dimension 1) We first investigate the orthogonal maps $\mathcal{A} : V \to V$ with $\dim(V) = 1$. All linear maps in a vector space of dimension 1 are of the form: muliplication with a scalar α (check this). The length is only invariant for $\alpha = 1$ or $\alpha = -1$. When $\alpha = 1$ then $\mathcal{A} = I$; directly orthogonal. When $\alpha = -1$ then $\mathcal{A}\underline{x} = -\underline{x}$; reflection in the origin.

2.1.25 (Dimension 2) Next we look at orthogonal maps $\mathcal{A} : V \to V$ with $\dim(V) = 2$. Choose an orthonormal basis $\alpha = \{\underline{a}_1, \underline{a}_2\}$. $\mathcal{A}\underline{a}_1$ in $\mathbb{R}^2$ has length 1 so can be written as $(\cos\varphi, \sin\varphi)$ for some ϕ. $\mathcal{A}\underline{a}_2$ also has length 1 and is perpendicular to $\mathcal{A}\underline{a}_1$, so for $\mathcal{A}\underline{a}_2$ we have two possibilities: $(-\sin\varphi, \cos\varphi)$ or $(\sin\varphi, -\cos\varphi)$.



In the first case we get the matrix

$$A_\alpha = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

This is the matrix of a *rotation over an angle* φ; this is easy to see by determining the angle between $\underline{x} = \lambda\underline{a}_1 + \mu\underline{a}_2$ and its image $\mathcal{A}\underline{x} = (\lambda\cos\varphi - \mu\sin\varphi)\underline{a}_1 + (\lambda\sin\varphi + \mu\cos\varphi)\underline{a}_2$, check this. The determinant is $+1$; the rotation is directly orthogonal.

In the second case we get the matrix

$$A_\alpha = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}.$$

This matrix has determinant -1. That means because of 2.1.23 that there is an eigenvalue -1 and sinces $\lambda_1{\cdot}\lambda_2 = -1$ the other eigenvalue is $+1$. So we have a basis $\beta = \{\underline{b}_1, \underline{b}_2\}$ of eigenvectors for eigenvalues -1 and 1 and the matrix w.r.t. this basis is

$$A_\beta = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The eigenvectors $\underline{b}_1$ and $\underline{b}_2$ are perpendicular, because

$$(\underline{b}_1, \underline{b}_2) = (\mathcal{A}\underline{b}_1, \mathcal{A}\underline{b}_2) = (-\,\underline{b}_1, \underline{b}_2) = -\,(\underline{b}_1, \underline{b}_2)$$

hence $(\underline{b}_1, \underline{b}_2) = 0$. From the matrix A_β we can read the geometric meaning of the map $\mathcal{A} : V \to V$. It is a reflection in the line $< \underline{b}_2 >$: a vector $\underline{x} = \lambda_1\underline{b}_1 + \lambda_2\underline{b}_2$ with components $\lambda_1\underline{b}_1$ perpendicular to $< \underline{b}_2 >$ and $\lambda_2\underline{b}_2$ in the direction $< \underline{b}_2 >$ is mapped to $-\lambda_1\underline{b}_1 + \lambda_2\underline{b}_2$.

To understand the structure of an orthogonal map $\mathcal{A} : V \to V$ in general, we need the following result concerning invariant subspaces.

2.1.26 Theorem. *Let W be a linear subspace of a real finite dimensional inner product space V and let $\mathcal{A} : V \to V$ be an orthogonal map such that W is invariant under $\mathcal{A}$. Then also $W^\perp$ is invariant under $\mathcal{A}$.*

Proof. Since $\mathcal{A}(W) \subseteq W$ and since $\mathcal{A}$ is injective, we have $\mathcal{A}(W) = W$. Let $\underline{v}$ be an arbitrary vector in $W^\perp$. For every vector $\underline{w} \in W$ there exists a vector $\underline{w}^* \in W$ with $\mathcal{A}\underline{w}^* = \underline{w}$, since $\mathcal{A}(W) = W$. Then

$$(A\underline{v}, \underline{w}) = (\mathcal{A}\underline{v}, \mathcal{A}\underline{w}^*) = (\underline{v}, \underline{w}^*) = 0,$$

since $\underline{v} \in W^\perp$ and $\underline{w}^* \in W$. The vector $\mathcal{A}\underline{v}$ is therefore perpendicular to all vectors of W, so if $\underline{v} \in W^\perp$ then also $\mathcal{A}\underline{v} \in W^\perp$; hence $W^\perp$ is invariant. $\square$

With Theorem 1.5.6 this leads to the following result:

2.1.27 Theorem. *Let $\mathcal{A}$ be an orthogonal map in a finite dimensional inner product space V and suppose W is an invariant linear subspace of V.Then there exists an orthonormal basis α such that*

$$A_\alpha = \begin{pmatrix} M_1 & O_1 \\ O_2 & M_2 \end{pmatrix},$$

where M_1 and M_2 are orthogonal matrices of the restrictions $\mathcal{A} : W \to W$ and $\mathcal{A} : W^\perp \to W^\perp$ respectively and O_1, O_2 zero matrices of the right size. In addition $\det(A_\alpha) = \det(M_1) \cdot \det(M_2)$.

2.1.28 (Dimension 3) We start with the case of a directly orthogonal map $\mathcal{A}$, that is with determinant 1. The product of the roots of the characteristic equation is then 1. If all zeroes of the characteristic polynomial are real (so equal to ± 1), then at least one of them is equal to 1. If there is a non-real root λ, then also $\bar{\lambda}$ is a root and now the third root has to be real and therefore 1. So there always is an eigenvalue 1. Let $\underline{a}_1$ be an eigenvector of length 1 with eigenvalue 1. Since $< \underline{a}_1 >$ is invariant, also the two-dimensional subspace $W =< \underline{a}_1 >^\perp$ is invariant. The restriction of $\mathcal{A}$ to W is again an orthogonal map with determinant 1 (because of Theorem 2.1.27). So there is an orthonormal basis $\{\underline{a}_2, \underline{a}_3\}$ of W with the property that the matrix of $\mathcal{A} : W \to W$ w.r.t. $\{\underline{a}_2, \underline{a}_3\}$ looks like

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

and the matrix of $\mathcal{A}$ w.r.t. the basis $\{\underline{a}_1, \underline{a}_2, \underline{a}_3\}$ like:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}.$$

The map is a *rotation* around the *axis (of rotation)* $< \underline{a}_1 >$ over an angle φ. Notice that the trace of the map equals $1 + 2\cos\varphi$; knowing the trace, for example from any matrix for $\mathcal{A}$ w.r.t. some basis, the angle can be quickly computed.

Now the case that the determinant is -1. In this case there is an eigenvalue -1. Let $\underline{a}_1$ be an eigenvector of length 1 for this eigenvalue -1. As above we see that $W =< \underline{a}_1 >^\perp$ is invariant and the restriction $\mathcal{A} : W \to W$ has determinant 1. W.r.t. a suitable orthonormal basis $\{\underline{a}_1, \underline{a}_2, \underline{a}_3\}$, $\mathcal{A}$ has matrix

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}.$$

The map is a so called *rotoreflection*: a rotation around the axis $< \underline{a}_1 >$ combined with a reflection in the plane W. Special cases arise if $\varphi = 0$ or $\varphi = \pi$. In the first case $\mathcal{A}$ is a reflection with *mirrorplane* W; in the second case $\mathcal{A} = -\mathcal{I}$, that is, every vector is mapped to its opposite.

2.1.29 The general case roughly goes as follows. Choose orthonormal bases of the eigenspaces E_1 and E_{-1} for the map $\mathcal{A}$. Together these bases span an invariant subspace W. So $W^\perp$ is also invariant and the characteristic polynomial of the restriction of $\mathcal{A} : W^\perp \to W^\perp$ only has non-real roots. Each of these roots corresponds according to 1.5.8 a two-dimensional invariant subspace. According to our analysis of the two-dimensional case the restriction of $\mathcal{A}$ to such a subspace is a rotation. With respect to a suitable orthonormal basis we therefore find a matrix of the following form.

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & 0 & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \boxed{D_1} & & \\ & & & 0 & & & & \ddots & \\ & & & & & & & & \boxed{D_r} \end{pmatrix}$$

here $D_1, \ldots, D_r$ are 2×2 rotation matrices.

2.1.30 **Example.** Consider the map $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^3$ determined by the matrix

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

This matrix is orthogonal so the map $\mathcal{A}$ is orthogonal as well. We first compute the determinant of A and find 1. $\mathcal{A}$ is therefore a rotation around the axis in the direction of an eigenvector with eigenvalue 1. This we find by solving $(\mathcal{A} - I)\underline{x} = \underline{0}$, so from

$$\frac{1}{3}\begin{pmatrix} -1 & 1 & -2 \\ 1 & -1 & 2 \\ 2 & -2 & -1 \end{pmatrix} \approx \begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

from which we compute $E_1 =< (1,1,0) >$. So this is the axis of rotation. The angle of rotation we find using the trace: $1 + 2\cos\varphi = \frac{5}{3}$, so $\cos\varphi = \frac{1}{3}$, $\varphi = \arccos(\frac{1}{3})$.

Another way to find this angle is as follows: the plane perpendicular to the axis has equation $x + y = 0$. Choose a vector in this plane, for example $\underline{a} = (1,-1,0)$. Then $\mathcal{A}\underline{a} = \frac{1}{3}(1,-1,4)$. This vector is of course again contained in the plane $x+y=0$, because this plane is invariant under $\mathcal{A}$. The angle of rotation is now the angle between $\underline{a}$ and $\mathcal{A}\underline{a}$. This we find using the inner product:

$$(\underline{a}, \mathcal{A}\underline{a}) = \|\underline{a}\|\,\|\mathcal{A}\underline{a}\|\cos\varphi$$

$$\tfrac{2}{3} = \sqrt{2}\,\sqrt{2}\,\cos\varphi,$$

so $\cos\varphi = (1/3)$. Geometrically is it clear that it is irrelevant which vector in the plane $x+y=0$ you take, but this also follows from the computation.

2.2 Symmetric maps

2.2.1 In this section we discuss a second important class of linear maps in a real inner product space V, the class of symmetric maps. An important role will be played by symmetric matrices. The following topics will be discussed

- the notion of a symmetric map and the connection with symmetric matrices,
- the main result, that symmetric linear maps always have a basis of eigenvectors,
- the notion of a quadratic form and the classification of quadratic forms, an important application of the theory of symmetric linear maps.

2.2.2 **Definition. (Symmetric map)** Let V be a real inner product space. The linear map $\mathcal{A} : V \to V$ is called *symmetric* if $(\mathcal{A}\underline{x}, \underline{y}) = (\underline{x}, \mathcal{A}\underline{y})$ for all $\underline{x}, \underline{y}$ in V.

2.2.3 **Example.** The orthogonal projection $\mathcal{P}$ on a line l is symmetric. If $l =< \underline{a} >$ and $\|\underline{a}\| = 1$, then $\mathcal{P}\underline{x} = (\underline{x}, \underline{a})\underline{a}$. For all $\underline{x}$ and $\underline{y}$ we now have

$$\begin{aligned}(\mathcal{P}\underline{x}, \underline{y}) &= ((\underline{x}, \underline{a})\underline{a}, \underline{y}) = (\underline{x}, \underline{a})(\underline{a}, \underline{y}),\\ (\underline{x}, \mathcal{P}\underline{y}) &= (\underline{x}, (\underline{y}, \underline{a})\underline{a}) = (\underline{y}, \underline{a})(\underline{x}, \underline{a}),\end{aligned}$$

so that $(\mathcal{P}\underline{x}, \underline{y}) = (\underline{x}, \mathcal{P}\underline{y})$.

In a similar way one can show that the orthogonal projection on a subspace W is a symmetric linear map.

2.2.4 The next theorem allows us make the connection with matrices.

2.2.5 Theorem. *For a linear map $\mathcal{A} : V \to V$ the following are equivalent:*

1. *$\mathcal{A} : V \to V$ is symmetric.*
2. *For every orthonormal system $\underline{a}_1, \ldots, \underline{a}_m$ in V we have $(\mathcal{A}\underline{a}_i, \underline{a}_j) = (\underline{a}_i, \mathcal{A}\underline{a}_j)$ for all i, j.*
3. *There is an orthonormal basis $\{\underline{a}_1, \ldots, \underline{a}_n\}$ of V satisfying $(\mathcal{A}\underline{a}_i, \underline{a}_j) = (\underline{a}_i, \mathcal{A}\underline{a}_j)$ for all i, j.*

Proof. 1) $\Rightarrow$ 2) and 2) $\Rightarrow$ 3) are trivial.

3) $\Rightarrow$ 1): Write $\underline{x} = \lambda_1\underline{a}_1 + \cdots + \lambda_n\underline{a}_n$ and $\underline{y} = \mu_1\underline{a}_1 + \cdots + \mu_n\underline{a}_n$. Then

$$(\mathcal{A}\underline{x}, \underline{y}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\mu_j(\mathcal{A}\underline{a}_i, \underline{a}_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\mu_j(\underline{a}_i, \mathcal{A}\underline{a}_j) = (\underline{x}, \mathcal{A}\underline{y}).$$

$\square$

2.2.6 (Connection with symmetric matrices) The name "symmetric" has to do with the matrix representation of $\mathcal{A}$ in the finite dimensional case. Let $\alpha : \{\underline{a}_1, \ldots, \underline{a}_n\}$ be an arbitrary *orthonormal* basis for V. The matrix A_α has as columns the α–coordinates of the vectors $\mathcal{A}\underline{a}_1, \mathcal{A}\underline{a}_2, \ldots, \mathcal{A}\underline{a}_n$. On the ij–th position we find i–th coordinate of $\mathcal{A}\underline{a}_j$, so $(\underline{a}_i, \mathcal{A}\underline{a}_j)$, and on the ji–th position the j–th coordinate of $\mathcal{A}\underline{a}_i$, so $(\mathcal{A}\underline{a}_i, \underline{a}_j)$. If $\mathcal{A}$ is symmetric, then these elements are equal, so the matrix is symmetric in the main diagonal. Formally: $A_\alpha^\top = A_\alpha$. Conversely, if A_α is symmetric, then the element on the ij–th position is equal to the element on the ji–th position, that is $(\underline{a}_i, \mathcal{A}\underline{a}_j) = (\underline{a}_i, \mathcal{A}\underline{a}_j)$. The above theorem shows then that $\mathcal{A}$ is symmetric. Hence:

2.2.7 Theorem. *Let V be a finite dimensional real inner product space and α an orthonormal basis of V. The linear map $\mathcal{A} : V \to V$ is symmetric if and only if the matrix A_α of $\mathcal{A}$ w.r.t. α is symmetric.*

2.2.8 Example. We can use this theorem to show (again) that orthogonal projection $\mathcal{P}_W : V \to V$ on a subspace W is symmetric, (see example 2.2.3). Choose an orthonormal basis $\{\underline{a}_1, \ldots, \underline{a}_n\}$ of V in such a way that $\{\underline{a}_1, \ldots, \underline{a}_m\}$ is an orthonormal basis of W. Every vector from the basis is an eigenvector: the first m vectors for eigenvalue 1, the last $n-m$ for eigenvalue 0. The matrix of $\mathcal{P}_W$ w.r.t. this orthonormal basis is therefore a diagonal matrix, hence symmetric. We now use the theorem to conclude that $\mathcal{P}_W$ is symmetric.

2.2.9 (Bases of eigenvectors) The following property – a similar property we have seen for orthogonal maps – is one of the two pillars on which the diagonalisability of symmetric maps rests. The other pillar is the theorem that comes immediately after it.

2.2.10 Theorem. *Let V be a real inner product space, $\mathcal{A} : V \to V$ a symmetric map and suppose W is an invariant linear subspace of V. Then also $W^\perp$ is invariant.*

Proof. Take an arbitrary vector $\underline{x} \in W^\perp$. We will show that $\mathcal{A}\underline{x} \in W^\perp$ by proving $(\mathcal{A}\underline{x}, \underline{y}) = 0$ for all $\underline{y} \in W$. This is done as follows: $(\mathcal{A}\underline{x}, \underline{y}) = (\underline{x}, \mathcal{A}\underline{y})$ because $\mathcal{A}$ is symmetric. Since $\underline{y} \in W$ and since W is invariant under $\mathcal{A}$ it follows that $\mathcal{A}\underline{y} \in W$. Since $\underline{x} \in W^\perp$ one has $(\underline{x}, \mathcal{A}\underline{y}) = 0$, so $(\mathcal{A}\underline{x}, \underline{y}) = 0$. □

2.2.11 Theorem. *Let $\mathcal{A} : V \to V$ be a symmetric linear map with* $\dim(V) < \infty$. *Then all roots of the characteristic equation are real.*

Proof. Suppose μ is a non-real root of the characteristic equation. According to 1.5.8 there is a 2-dimensional invariant linear subspace W such that the restriction $\mathcal{A} : W \to W$ has a characteristic equation with roots μ and $\bar{\mu}$, (since μ is non-real). Choose an orthonormal basis α for W. Then

$$A_\alpha = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

is symmetric according to Theorem 2.2.7, with characteristic polynomial

$$\begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = \lambda^2 - (a+c)\lambda + ac - b^2.$$

The discriminant of this quadratic is $(a+c)^2 - 4ac + 4b^2 = (a-c)^2 + 4b^2 \geq 0$, so the two roots are real. This is a contradiction, so all roots of the characteristic equation of $\mathcal{A} : V \to V$ are real. □

2.2.12 Theorem. *Let $\mathcal{A} : V \to V$ be a symmetric linear map, with* $\dim(V) < \infty$*. Then there exists an orthonormal basis of eigenvectors of $\mathcal{A}$.*

Proof. Let $\dim(V) = n$. The characteristic equation of $\mathcal{A} : V \to V$ has only real roots, so there are eigenvalues. Choose an eigenvalue λ_1 and let $\underline{v}_1$ be an eigenvector with $\|\underline{v}_1\| = 1$. Let $W_1 =< \underline{v}_1 >^\perp$. Then W_1 is invariant, $\dim(W_1) = n - 1$, the restriction $\mathcal{A} : W_1 \to W_1$ is again symmetric so the characteristic equation of $\mathcal{A} : W_1 \to W_1$ has real roots only, and therefore eigenvalues. Let $\underline{v}_2$ be an eigenvector in W_1 for eigenvalue λ_2 and $\|\underline{v}_2\| = 1$. Then $\underline{v}_1, \underline{v}_2$ is an orthonormal system and $< \underline{v}_1, \underline{v}_2 >$ is invariant. Also $W_2 =< \underline{v}_1, \underline{v}_2 >^\perp$ is invariant, $\dim W_2 = n - 2$, and the restriction $\mathcal{A} : W_2 \to W_2$ is also symmetric, so the characteristic equation of $\mathcal{A} : W_2 \to W_2$ has real roots only, so eigenvalues. We choose an eigenvector $\underline{v}_3 \in W_2$ with eigenvalue λ_3 and $\|\underline{v}_3\| = 1$. Then $\underline{v}_1, \underline{v}_2, \underline{v}_3$ is an orthonormal system, etc. In this way we find in the end an orthonormal basis of eigenvectors of V. □

2.2.13 It follows from this construction that eigenvectors belonging to different eigenvalues are mutually perpendicular. This can also be seen directly. Consider the eigenspaces E_λ and E_μ with $\lambda \neq \mu$. Take $\underline{x} \in E_\lambda$ and $\underline{y} \in E_\mu$. Then

$$\lambda(\underline{x}, \underline{y}) = (\lambda\underline{x}, \underline{y}) = (\mathcal{A}\underline{x}, \underline{y}) = (\underline{x}, \mathcal{A}\underline{y}) = (\underline{x}, \mu\underline{y}) = \mu(\underline{x}, \underline{y}),$$

whence $(\underline{x}, \underline{y}) = 0$ and so $E_\lambda \perp E_\mu$.

2.2.14 Corollary. *A symmetric matrix is diagonalisable by changing to an orthonormal basis of eigenvectors, so by means of an orthogonal coordinate transformation and hence an orthogonal transition matrix.*

2.2.15 Example. For the symmetric matrix

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

we find eigenvalues $0, -4$ and 6 with eigenspaces

$$E_0 =< (-1, 1, 1) >, \ E_{-4} =< (1, 0, 1) >, \ E_6 =< (1, 2, -1) > .$$

An orthonormal basis of eigenvectors is therefore

$$\alpha = \left\{ \frac{1}{\sqrt{3}}(-1, 1, 1), \ \frac{1}{\sqrt{2}}(1, 0, 1), \ \frac{1}{\sqrt{6}}(1, 2, -1) \right\}.$$

The matrix of A_α can be written down without further computations, because on the diagonal the eigenvalues appear:

$$A_\alpha = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

2.2.16 (Application: quadratic forms) As an application of symmetric matrices we show how to bring quadratic curves and surfaces "on principal axes".

It is easy to verify that

$$(x_1 \ \cdots \ x_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2$$

$$+(a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + \cdots + (a_{n-1,n} + a_{n,n-1})x_{n-1}x_n \ .$$

So the result is a homogeneous polynomial of degree 2 in $x_1, \dots, x_n$. The coefficients of $x_1^2, \dots, x_n^2$ are the diagonal elements of the matrix and the coefficient of x_ix_j with $i \neq j$ is the sum of a_{ij} and a_{ji}. Going in the other direction, this result means that we can write every quadratic form in $x_1, \dots, x_n$ (that is, a polynomial in $x_1, \dots x_n$ with all terms of degree 2) as a matrix product, and even with a *symmetric* matrix A. Diagonal elements of A are coefficients of $x_1^2, \dots, x_n^2$ and for $i \neq j$ we take $a_{ij} = a_{ji}$ is half of the coefficient of x_ix_j.

2.2.17 Example.

$$2x^2 - 4xy + 3y^2 = (x \ y) \begin{pmatrix} 2 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

$$2x^2 - 4xy + 4xz - 3y^2 - 3z^2 = (x \ y \ z) \begin{pmatrix} 2 & -2 & 2 \\ -2 & -3 & 0 \\ 2 & 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Check this!

2.2.18 Consider next a quadratic form

$$\varphi(x_1, \dots, x_n) = (x_1 \ \dots \ x_n)A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

for some symmetric matrix A. Let $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ be an orthonormal basis of eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$. We denote the coordinates of $(x_1, \ldots, x_n)$ w.r.t. the basis α by $(y_1, \ldots, y_n)$. So the connection is

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = {}_\varepsilon S_\alpha \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

where the columns of the matrix ${}_\varepsilon S_\alpha$ are the vectors $\underline{a}_1, \ldots, \underline{a}_n$. After transposition this relation reads $(x_1, \ldots, x_n) = (y_1, \ldots, y_n) {}_\varepsilon S_\alpha^\top$. Since the matrix ${}_\varepsilon S_\alpha$ is orthogonal ${}_\varepsilon S_\alpha^\top = ({}_\varepsilon S_\alpha)^{-1}$ and this matrix in turn equals ${}_\alpha S_\varepsilon$. Substituting this in the quadratic form changes this into

$$(y_1 \ \ldots \ y_n) \ {}_\alpha S_\varepsilon \ A \ {}_\varepsilon S_\alpha \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= (y_1 \ \ldots \ y_n) \ A_\alpha \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

because A_α is a diagonal matrix is with on the diagonal the eigenvalues. After this coordinate transformation dubble products have disappeared! Notice that we can write down this sum of squares as soon as we know the eigenvalues (including their multiplicities).

2.2.19 Example. Consider in the plane the curve

$$x^2 - 4xy + y^2 = 2$$

where (x, y) are the coordinates of a vector w.r.t. an orthonormal basis. We can write the left hand side as

$$(x \ \ y) \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} .$$

The eigenvalues of this matrix are 3 and -1, with eigenspaces $E_3 =< (1, -1) >$, $E_{-1} =< (1, 1) >$. So an orthonormal basis of eigenvectors is $\alpha = \{\frac{1}{2}\sqrt{2}(1, -1), \frac{1}{2}\sqrt{2}(1, 1)\}$. It follows that

$${}_\varepsilon S_\alpha = \tfrac{1}{2}\sqrt{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

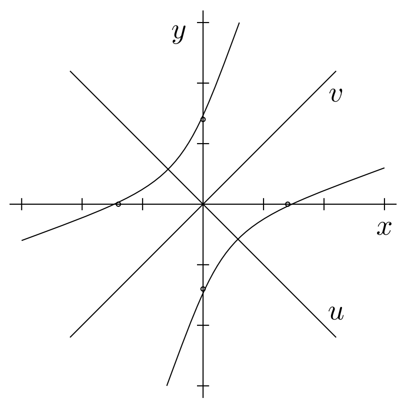
so

$$\begin{pmatrix} x \\ y \end{pmatrix} = \tfrac{1}{2}\sqrt{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

When we substitute this relation in the equation of the curve we get (without computation!)

$$3u^2 - v^2 = 2.$$

So the curve is a hyperbola and the principal axes are in the direction of the eigenvectors $(1,-1)$ and $(1,1)$.



2.2.20 Quadratic curves and surfaces in general also contain linear terms. Typically they can be removed by a translation of the coordinate system (so choosing a different origin). We first indicate how to do this, and then discuss an example.

A quadratic (hyper)surface can be described as

$$(x_1 \ \ldots \ x_n)\ A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + (b_1 \ \ldots \ b_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + d = 0;$$

here A is a symmetric matrix. Let $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ be an orthonormal basis of eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$. We denote the coordinates of $(x_1, \ldots, x_n)$ w.r.t. α by $(y_1, \ldots, y_n)$. So the relation is

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = {}_{\varepsilon}S_{\alpha} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

where the columns of ${}_\varepsilon S_\alpha$ are of course the vectors $\underline{a}_1, \ldots, \underline{a}_n$. After substitution in the equation:

$$(y_1 \ \ldots \ y_n)\ {}_\varepsilon S_\alpha^\top \ A \ {}_\varepsilon S_\alpha \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + (b_1 \ \ldots \ b_n)\ {}_\varepsilon S_\alpha \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + d = 0 \ ,$$

$$\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 + c_1 y_1 + \cdots + c_n y_n + d = 0 \ .$$

Here

$$(c_1 \ \ldots \ c_n) = (b_1 \ \ldots \ b_n)\ {}_\varepsilon S_\alpha \,, \text{ hence } \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = {}_\varepsilon S_\alpha^\top \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \ .$$

$(c_1, \ldots, c_n)$ are the the α–coordinates of the vector $\underline{b}$.

If $\lambda_i \neq 0$, then we write

$$\lambda_i y_i^2 + c_i y_i = \lambda_i \Big(y_i^2 + \frac{c_i}{\lambda_i} y_i \Big) \ = \lambda_i \Big(y_i + \frac{c_i}{2\lambda_i} \Big)^2 - \frac{c_i^2}{4\lambda_i} \ .$$

For all i with $\lambda_i \neq 0$ we now substitute $y_i + \frac{c_i}{2\lambda_i} = z_i$. The result is a polynomial in which each of the variables (z_i) occurs in exactly one term, linear of quadratic.

2.2.21 Example. We consider the quadratic curve with equation

$$16x^2 - 24xy + 9y^2 - 30x - 40y = 0,$$

$$(x \ \ y) \begin{pmatrix} 16 & -12 \\ -12 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (-30 \ \ -40) \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

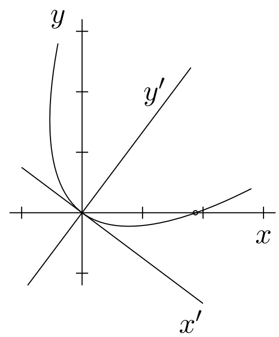
The eigenvalues of the matrix are 25 and 0 with eigenspaces $E_{25} =< (4, -3) >$ and $E_0 =< (3, 4) >$. We take the orthonormal basis $\{ \frac{1}{5}(4, -3), \ \frac{1}{5}(3, 4) \ \}$. The relation between the coordinates is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} .$$

After substitution we get

$$25x'^2 + 0y'^2 - 30 \cdot \tfrac{1}{5}(4x' + 3y') - 40 \cdot \tfrac{1}{5}(-3x' + 4y') = 0,$$

which reduces to $25x'^2 - 50y' = 0$. The curve is a parabola, w.r.t. the basis α the equation is $y' = \frac{1}{2}x'^2$.



2.2.22 Example. We consider in the plane the quadratic curve

$$91x^2 - 24xy + 84y^2 - 770x - 360y + 2200 = 0 \ .$$

The matrix of the quadratic part is

$$\begin{pmatrix} 91 & -12 \\ -12 & 84 \end{pmatrix}.$$

The eigenvalues are 75 and 100 with eigenspaces $E_{75} =< (3,4) >$, $E_{100} =< (-4,3) >$. An orthonormal basis of eigenvectors is then $\alpha = \{(\frac{1}{5}(3,4), \frac{1}{5}(-4,3)\}$. Let (x', y') be the coordinates of (x, y) w.r.t. α. Then

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

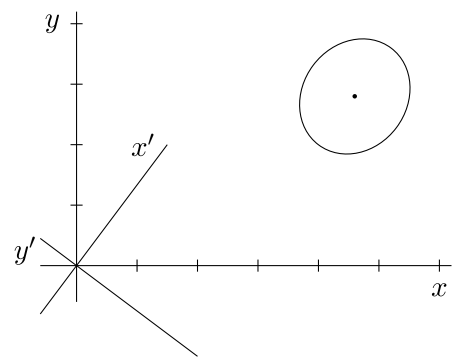
When we substitute this in the equation of the curve we get

$$75x'^2 + 100y'^2 - 750x' + 400y' + 2200 = 0,$$

$$3x'^2 + 4y'^2 - 30x' + 16y' + 88 = 0,$$

$$3(x' - 5)^2 + 4(y' + 2)^2 = 3.$$

We see that the quadratic curve is an ellipse. The α–coordinates of the centre are $(5, -2)$; the centre therefore is $(\frac{23}{5}, \frac{14}{5})$. The directions of the axes are (in x, y–coordinates) $(3, 4)$ and $(-4, 3)$.



2.3 Notes

The notion orthogonal matrix originates in the work of G.F. Frobenius (1849–1917). Orthogonality relations occur and are important in his work on group representations. Diagonalisability of symmetric matrices and the application to the study of quadratic forms are found in the works of Cauchy.

Algebra

Quadratic forms occur in analysis in the investigation of the behaviour of a function in the neighborhood of a stationary punt. Second-order approximations tell us whether the function has a local minimum or maximum, or a saddle point or worse.

Analyse in D

2.4 Exercises

§1

1 The linear map $\mathcal{A}\colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by the matrix

$$A = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}.$$

Show that $\mathcal{A}$ is an orthogonal map. Is $\mathcal{A}$ a rotation or a reflection. In the first case, determine the angle of rotation, in the second case the mirror-line.

2 The linear map $\mathcal{A}\colon \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\mathcal{A}\underline{x} = \underline{x} + \lambda(\underline{a}, \underline{x})\underline{a}, \quad \underline{a} \in \mathbb{R}^n,\ ||\underline{a}|| = 1,\ \lambda \in \mathbb{R}.$$

a. For what values of λ is the map $\mathcal{A}$ orthogonal?

b. Give an orthonormal basis α of $\mathbb{R}^n$ w.r.t. which the matrix A_α is diagonal.

3 Let $\alpha = \{\underline{a}_1, \underline{a}_2, \underline{a}_3, \underline{a}_4\}$ be an orthonormal basis of $\mathbb{R}^4$. The linear map $\mathcal{A}\colon \mathbb{R}^4 \to \mathbb{R}^4$ has the following properties:

$$\mathcal{A}^2 = \mathcal{I}, \quad \mathcal{A}\underline{a}_1 = \underline{a}_3, \quad \mathcal{A}\underline{a}_2 = \underline{a}_4.$$

a. Show that $\mathcal{A}$ is orthogonal.

b. Determine the eigenvalues and an orthonormal basis of eigenvectors of $\mathcal{A}$.

4 We know the following about the linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$:

a. $\mathcal{A}$ is orthogonal,

b. $\mathcal{A}(1,2,2) = (1,2,2)$,

c. The vector $(2,0,-1)$ is eigenvector for eigenvalue -1,

d. $\dim E_1 = 1$.

Determine the matrix of $\mathcal{A}$.

5 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the rotation around $< (1,0,1) >$ over an angle $\arccos(3/5)$ to the right (from the direction $(1,0,1)$).

a. Determine an orthonormal basis α of $\mathbb{R}^3$ containing the vector $\frac{1}{\sqrt{2}}(1,0,1)$. What does the matrix A_α look like?

b. Determine the transition matrices ${}_\varepsilon S_\alpha$ and ${}_\alpha S_\varepsilon$ and then determine the matrix of $\mathcal{A}$.

6 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

$$A = \frac{1}{12}\begin{pmatrix} 5 & -7 & 5 \\ 2 & 2 & -10 \\ -7 & 5 & 5 \end{pmatrix}.$$

a. Show that the range of $\mathcal{A}$ is the plane V with equation $x_1 + x_2 + x_3 = 0$. Explain why $\mathcal{A}$ is *not* orthogonal.

b. The plane V from part a.) is an invariant subspace. Determine an orthonormal basis α of V and the 2×2 matrix of the restriction $\mathcal{A}\colon V \to V$ w.r.t. α.

c. Show that the restriction $\mathcal{A}\colon V \to V$ is an orthogonal map. What is the geometric meaning of this map?

d. We know for the linear map $\mathcal{B}\colon \mathbb{R}^3 \to \mathbb{R}^3$ that $\mathcal{B}\underline{x} = \mathcal{A}\underline{x}$ for all $\underline{x} \in V$ and that $\mathcal{B}$ is directly orthogonal. Determine the matrix of $\mathcal{B}$.

7 The linear map $\mathcal{A}\colon \mathbb{R}^4 \to \mathbb{R}^4$ is given by the matrix

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

a. Show that the map $\mathcal{A}$ is orthogonal.

b. Determine the roots of the characteristic equation of $\mathcal{A}$. Show that $\mathbb{R}^4$ has no basis consisting of eigenvectors.

c. Determine all 1- and 2-dimensional invariant subspaces of $\mathcal{A}$.

d. Give an orthonormal basis α of $\mathbb{R}^4$ with the property that the matrix A_α only has 1×1 and 2×2 blocks with non-zero entries along the diagonal. Determine such a matrix A_α.

8 Each of the following matrices determines an orthogonal map on $\mathbb{R}^3$. Determine for each of these maps whether it is a rotation or a rotoreflection. Determine in case of a rotation the axis and angle of rotation and in case of a rotoreflection the mirror-plane ane the angle of rotation.

a. $\frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{pmatrix}$,

b. $\frac{1}{2}\begin{pmatrix} -1 & -1 & -\sqrt{2} \\ -1 & -1 & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} & 0 \end{pmatrix}$,

c. $\frac{1}{2}\begin{pmatrix} 1 & 1 & -\sqrt{2} \\ 1 & 1 & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} & 0 \end{pmatrix}$,

d. $\frac{1}{21}\begin{pmatrix} -5 & 4 & -20 \\ -20 & -5 & 4 \\ 4 & -20 & -5 \end{pmatrix}$.

9 The following matrices describe orthogonal maps from $\mathbb{R}^3$ to $\mathbb{R}^3$. Determine in each case an orthonormal basis α of $\mathbb{R}^3$ w.r.t. which the matrix of the map is of the form

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

where $\lambda = \pm 1$ and $0 \leq \varphi \leq \pi$ and give λ and φ.

a. $\begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, b. $\begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

§2

10 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$\mathcal{A}\underline{x} = 2\underline{x} - (\underline{a}, \underline{x})\underline{a}, \quad \underline{a} \in \mathbb{R}^3,\ ||\underline{a}|| = 1.$$

a. Show that the map $\mathcal{A}$ is symmetric.

b. Show that $\underline{a}$ is an eigenvector. Determine a basis of eigenvectors for $\mathcal{A}$ and give the matrix of $\mathcal{A}$ w.r.t. deze basis.

11 We provide $\mathbb{R}^3$ with the basis $\alpha = \{\underline{a}_1, \underline{a}_2, \underline{a}_3\}$ where

$$\underline{a}_1 = (1,0,0), \quad \underline{a}_2 = (-1,1,0), \quad \underline{a}_3 = (0,0,1).$$

The matrix A_α of the linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is

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

a. show that the basis α is not orthonormal.

b. Show that the map $\mathcal{A}$ is symmetric, despite the fact that the matrix A_α is not.

c. Determine the eigenvalues and the corresponding eigenspaces of $\mathcal{A}$.

12 The linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$\begin{aligned} \mathcal{A}(2,-1,2) &= (0,0,0), \\ \mathcal{A}(-1,2,2) &= (-1,2,2), \\ \mathcal{A}(2,2,-1) &= (2,2,-1). \end{aligned}$$

a. Determine the null space and the range of $\mathcal{A}$.

b. Show that the map $\mathcal{A}$ is symmetric.

c. What is the geometric description of the map $\mathcal{A}$?

13 Concerning the symmetric linear map $\mathcal{A}\colon \mathbb{R}^3 \to \mathbb{R}^3$ we know the following:

i) the range of $\mathcal{A}$ is $\{\underline{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$,

ii) $\mathcal{A}(1,-1,0) = (1,-1,0)$,

iii) the characteristic roots $\lambda_1, \lambda_2, \lambda_3$ of $\mathcal{A}$ satisfy $\lambda_1 + \lambda_2 + \lambda_3 = 0$.

Determine the matrix of $\mathcal{A}$.

14 Let V be a real inner product space. If $\mathcal{A}\colon V \to V$ is a symmetric linear map, then the null space of $\mathcal{A}$ is the orthoplement of the range of $\mathcal{A}$. Show this.

15 Let l be a line in $\mathbb{R}^2$ through $(0,0)$, and let $\mathcal{P}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the orthogonal projection on l and let $\mathcal{S}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the orthogonal reflection in l. Show that the linear maps $\mathcal{P}$ and $\mathcal{S}$ are symmetric by studying the matrix of these maps w.r.t. a suitable basis.

16 The following symmetric matrices correspond to symmetric linear maps from $\mathbb{R}^2$ to $\mathbb{R}^2$ or $\mathbb{R}^3$ to $\mathbb{R}^3$. Determine in each case an orthonormal basis of eigenvectors and the matrix of the map w.r.t. these bases.

a. $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, b. $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$,

c. $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$, d. $\begin{pmatrix} 1 & -4 & 8 \\ -4 & 7 & 4 \\ 8 & 4 & 1 \end{pmatrix}$.

17 Let V be a real inner product space and let $\mathcal{A}\colon V \to V$ and $\mathcal{B}\colon V \to V$ be linear maps. Prove the following statements:

a. If $\mathcal{A}$ and $\mathcal{B}$ are symmetric and $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, then also the product $\mathcal{A}\mathcal{B}$ is symmetric.

b. If $\mathcal{A}$ is symmetric and has an inverse, then the inverse $\mathcal{A}^{-1}$ is also symmetric.

18 Write the following curves in $\mathbb{R}^2$ on principal axes and determine the nature of the curve. Determine also the equations of the axes of symmetry in x_1, x_2–coordinates and finally give a sketch of the figure.

a. $x_1^2 + 8x_1x_2 + x_2^2 = 45$,

b. $3x_1^2 - 2x_1x_2 + 3x_2^2 = 2$,

c. $3x_1^2 + 8x_1x_2 - 3x_2^2 + 2x_1 - 14x_2 = 3$,

d. $4x_1^2 - 12x_1x_2 + 9x_2^2 - 7x_1 + 4x_2 = 12$,

e. $x_1^2 - 4x_1x_2 + 4x_2^2 - 30x_1 + 10x_2 + 25 = 0$.

19 In $\mathbb{R}^2$ we are given the ellipse with equation

$$7x_1^2 - 4x_1x_2 + 4x_2^2 - 6x_1 - 12x_2 = 9.$$

Determine the equation of the greatest circle that is contained in this ellipse and determine the two intersection points of this circle and the ellipse.

2.4.1 Exercises on exam level

20 The linear map $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^3$ is given by the matrix

$$A = \frac{1}{4}\begin{pmatrix} 3 & 1 & \sqrt{6} \\ 1 & 3 & -\sqrt{6} \\ -\sqrt{6} & \sqrt{6} & 2 \end{pmatrix}.$$

a. Prove that $\mathcal{A}$ is a rotation.

b. Determine the axis and angle of rotation of $\mathcal{A}$.

21 About the direct orthogonal map $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^3$ we know the following

$$\mathcal{A}(\underline{e}_1 + \underline{e}_2) = \underline{e}_1 + \underline{e}_2 \quad \text{and} \quad \mathcal{A}(\underline{e}_1 - \underline{e}_2) = -\sqrt{2}\,\underline{e}_3.$$

a. Determine the matrix A of $\mathcal{A}$ w.r.t. the standard basis.

b. The linear map $\mathcal{S} : \mathbb{R}^3 \to \mathbb{R}^3$ is the orthogonal reflection in the plane with equation $x_1 - x_2 = 0$. Determine the eigenvalues and an orthonormal basis of eigenvectors of $\mathcal{S}$.

c. show that the composite map $\mathcal{S} \circ \mathcal{A}$ is a reflection and determine the mirror plane of this map.

22 In $\mathbb{R}^2$ we are given the quadratic curve

$$2x_1^2 + 2x_1x_2 + 2x_2^2 + \sqrt{2}\,x_1 + 5\sqrt{2}\,x_2 = 2.$$

a. Determine the equation of this curve on principal axes. What is the nature of this curve?

b. Determine in (x_1, x_2)–coordinates the points of the curve with maximal distance to the centre

23

3. V is a 2-dimensional inner product space with an orthonormal basis $\alpha = \{\underline{a}, \underline{b}\}$. $\mathcal{B} : V \to V$ is the linear map with matrix w.r.t. the basis α

$$B_\alpha = \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix}.$$

The linear map $\mathcal{A} : V \to V$ is defined by

$$\mathcal{A}\underline{x} = 4\underline{x} - (\mathcal{B}\underline{x}, \underline{a})\underline{b} - (\mathcal{B}\underline{x}, \underline{b})\underline{a}.$$

a. Prove that the map $\mathcal{A}$ is symmetric.

b. Determine the eigenvalues and eigenvectors of $\mathcal{A}$.

Chapter 3

Decompositions

3.1 LU-decomposition

3.1.1 One of the central problems in linear algebra is the reduction of a problem to solving a system of linear equations $A\underline{x} = \underline{b}$. Once this reduction has been made, one still has to decide what strategy should be taken to solve this system. In case one has to solve very many such systems for a fixed (large) matrix A then it is worth while to compute the inverse of A. If on the other hand you have just one particular instance to solve, then Gaussian elimination so solving the system by bringing the matrix in row reduced form is probably the way to proceed. In practice different approaches are used, depending on the size of the matrix, how often the same matrix is used for a system of equations, and how sensitive the problem is for the accuracy of the coefficients. In this section will study one of these.

3.1.2 Suppose we have to solve the system $A\underline{x} = \underline{b}$, where for A we have a representation $A = LU$, with L a lower triangular matrix, so $L_{ij} = 0$ for $j > i$, and U an upper triangular matrix, so $U_{ij} = 0$ for $j < i$. This system can be solved very efficiently in two steps, we first solve $L\underline{y} = \underline{b}$ and then $U\underline{x} = \underline{y}$.

3.1.3 Example. We look at the following system $A\underline{x} = \underline{b}$ with $A = LU$:

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 11 & 8 \\ 8 & 19 & 26 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \underline{b} = \begin{pmatrix} 9 \\ 42 \\ 88 \end{pmatrix}.$$

First we solve $L\underline{y} = \underline{b}$ and find without any real computations after one another: $y_1 = 9$, $y_2 = (42 - 2 \cdot 9)/3 = 8$ and $y_3 = (88 - 4 \cdot 9 - 5 \cdot 8)/6 = 2$.

Next we solve $U\underline{x} = \underline{y}$, and find, again without any real computations, in the opposite order $x_3 = 2/2 = 1$, $x_2 = (8 - 2 \cdot 1)/3 = 2$ and finally $x_1 = (9 - 2 - 1)/2 = 3$.

3.1.4 Definition. (LU-decomposition) A factorisation of the square matrix A as $A = LU$ with L a lower triangular and U an upper triangular matrix is called an LU-decomposition of the matrix A.

3.1.5 Multiplication with a diagonal matrix doesn't change lower or upper triangularity, so the LU-decomposition is not uniquely determined. If a non-singular matrix A has an LU-decomposition then obviously L and U are non-singular, in particular we have that L_{11} and U_{11}, and hence A_{11} are non-zero, so we see that not every matrix has an LU-decomposition. In practice this is no problem, in a system of linear equations $A\underline{x} = \underline{b}$ we may freely change the order of the equations, which means permuting the rows of A.

3.1.6 If the matrix A can be reduced to (row reduced) normal form, without interchanging rows, then the matrix A itself has an LU-decomposition, we start by using row 1 to clear the remaining entries in the first column (by which we mean to turn the remaining entries in column 1 to zero), then row 2, to clear the entries beneath the diagonal in column 2 and so on. Each of these operations corresponds to left multiplication by lower triangular elementary matrices In this way we obtain $E_{n-1} \cdots E_1 A = U$, where E_i is the product of the elementaire matrices used to clear column i beneath the diagonal. U is upper triangular, all E_i and their inverses are lower triangular and we find $A = LU$ with $L = E_1^{-1} E_2^{-1} \cdots E_{n-1}^{-1}$.

3.1.7 Example. Determine an LU-decomposition of the matrix $\begin{pmatrix} 2 & 1 & 1 \\ 4 & 11 & 8 \\ 8 & 19 & 26 \end{pmatrix}$.

We first clear the first column with the matrix $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$ and get $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 9 & 6 \\ 0 & 15 & 22 \end{pmatrix}$. Now we clear the second column using the matrix $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{5}{3} & 1 \end{pmatrix}$ and find $U = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 9 & 6 \\ 0 & 0 & 12 \end{pmatrix}$. Finally we compute

$$L = E_1^{-1}E_2^{-1}\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{5}{3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & \frac{5}{3} & 1 \end{pmatrix}.$$

We don't actually have to compute the inverses of the elementary matrices and multiply them, because of their special form and order we essentially only have to do some bookkeeping:

3.1.8 Example. Determine an LU-decomposition of the matrix $\begin{pmatrix} 3 & 1 & 1 \\ 9 & 11 & 8 \\ 27 & 25 & 26 \end{pmatrix}$

Again we clear the columns systematically, but in stead of writing the 0 we created in the column we are working on, we write $\boxed{\text{c}}$ when we subtracted our 'working row' c times from the row we are working on:

$$\begin{pmatrix} 3 & 1 & 1 \\ 9 & 11 & 8 \\ 27 & 25 & 26 \end{pmatrix} \sim \begin{pmatrix} 3 & 1 & 1 \\ \boxed{3} & 8 & 5 \\ \boxed{9} & 16 & 17 \end{pmatrix} \sim \begin{pmatrix} 3 & 1 & 1 \\ \boxed{3} & 8 & 5 \\ \boxed{9} & \boxed{2} & 7 \end{pmatrix}.$$

Now U is the upper triangular part, with all boxes made empty: $U = \begin{pmatrix} 3 & 1 & 1 \\ 0 & 8 & 5 \\ 0 & 0 & 7 \end{pmatrix}$ and L is the matrix containing the. boxed entries, together with ones on the diagonal: $\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 9 & 2 & 1 \end{pmatrix}$.

We leave the proof of the correctness of the above procedure to the reader.

3.2 Singular Value Decomposition (SVD)

3.2.1 An important property of real symmetric matrices is that they have an orthonormal basis of eigenvectors (and so in particular they are diagonalisable). This has an interesting application when we are satisfied with a good approximation of our matrix, that is easier (cheaper) to store. If the symmetric matrix A has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with orthonormal basis of eigenvectors $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$, then we may write A as the sum of rank one matrices as follows:

$$A = \lambda_1 \underline{v}_1 \underline{v}_1^\top + \lambda_2 \underline{v}_2 \underline{v}_2^\top + \cdots + \lambda_n \underline{v}_n \underline{v}_n^\top.$$

To see this, multiply left and right with the vector $\underline{v}_i$, $i = 1, \ldots, n$.

In case that A has a few, say k large eigenvalues (large in absolute value), and many small ones, we can compute a good approximation of A by just taking these k terms in the sum above, the information we have to store is $k(n+1)$ numbers, namely λ_i and $\underline{v}_i$ for $i = 1, \ldots, k$, in stead of the n^2 coefficients of the matrix A. Here we assume we have ordered the eigenvalues in such a way that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$.

For non-square matrices we have no eigenvalues or eigenvectors and in this section we'll see what can be done in this case, because we still want to be able to compute a good but cheap approximation like we have for symmetric matrices.

3.2.2 Definition. (Singular value) If A is an $m \times n$ matrix, and the symmetric $n \times n$ matrix $A^\top A$ has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, then $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \ldots, n$ are the singular values of the matrix A.

3.2.3 Example. The matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ has singular values $\sqrt{3}$ and 1, because $A^\top A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ with eigenvalues 3 and 1. The transposed matrix $A^\top = B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ has singular values $\sqrt{3}, 1$ and 0 because $B^\top B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ with eigenvalues $3, 1$ and 0.

3.2.4 In the example we see that A and $A^\top$ have the same singular values $\neq 0$. Since $A^\top A$ is symmetric we know that A has an orthonormal basis of eigenvectors, and because $\underline{x}^\top A^\top A \underline{x} \geq 0$ for all (eigen)vectors, their eigenvalues are non-negative, which means that the singular values are well-defined. The singular values play a role a little bit similar to the eigenvalues of A (when present).

If A is symmetric, then clearly the singular values are precisely the absolute values of the eigenvalues of A. For a symmetric matrix A we have a decomposition $A = S^\top DS$, for an orthogonal matrix S, and a diagonal matrix D, with the eigenvalues on the diagonal.

The singular value decomposition (SVD) is a slightly weaker variation of this for general matrices.

3.2.5 Definition. (Generalised diagonal matrix, diagonal coefficients) An

$m \times n$ matrix A is called a generalised diagonal matrix if $A_{ij} = 0$ whenever $i \neq j$. The numbers A_{ii} are the diagonal coefficients of A.

3.2.6 Example. The matrix $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$ is a generalised diagonal matrix with diagonal coefficients 2 and 3.

3.2.7 Theorem. *Every $m \times n$ matrix A has a decomposition of the form $UDV^\top$, for an $m \times n$ generalised diagonal matrix D, whose diagonal coefficients are the singular values of A, and orthogonal $m \times m$ and $n \times n$ matrices U and V.*

3.2.8 The matrix A has n singular values, but the matrix D has m diagonal coefficients: The above theorem should be read as follows, the non-zero diagonal coefficients of D are precisely the non-zero singular values of A. It is clear that if $A = UDV^\top$ with orthogonal matrices U, V and a generalised diagonal matrix D, then $A^\top A = V(D^\top D)V^\top$, so the eigenvalues of $D^\top D$ are exactly those of $A^\top A$. So from this theorem it follows immediately that $A^\top$ and A have the same non-zero singular values, and since $A^\top = VD^\top U^\top$, they also have essentially the same singular value decomposition.

3.2.9 We now describe how to find the SVD $UDV^\top$ of an $m \times n$ matrix A. We start by computing an orthonormal basis $\{\underline{v}_1, \ldots, \underline{v}_n\}$ of eigenvectors of $A^\top A$, ordered in such a way that for the corresponding eigenvalues we have $\lambda_1 \geq \cdots \geq \lambda_n$. These will be the columns of the matrix V. The column vectors of the matrix U are computed with the formula $\underline{u}_j = \dfrac{1}{\sigma_j} A\underline{v}_j$, if $\sigma_j \neq 0$, and can be completed to an orthonormal basis if this does not yet define U. We finally find for A the desired formula $A = \sum_j \sigma_j \underline{u}_j \underline{v}_j^\top$ as the sum of rank one matrices that we can use for our approximation by deleting the small $\sigma_j's$.

3.2.10 Example. We determine a SVD for the matrix $A = \begin{pmatrix} 14 & 16 & 4 \\ -2 & -13 & -22 \end{pmatrix}$.

For this matrix we compute $A^\top A = \begin{pmatrix} 200 & 250 & 100 \\ 250 & 425 & 350 \\ 100 & 350 & 500 \end{pmatrix}$ with eigenvalues 900, 225 and 0 and an orthonormal basis of eigenvectors: $\underline{v}_1 = \frac{1}{3}(1, 2, 2)$, $\underline{v}_2 = \frac{1}{3}(2, 1, -2)$, $\underline{v}_3 = \frac{1}{3}(2, -2, 1)$. We find $\sigma_1 = 30$ and $\sigma_2 = 15$, and

$\underline{u}_1 = \dfrac{1}{30}A\underline{v}_1 = \dfrac{1}{5}(3,-4)$ and $\underline{u}_2 = \dfrac{1}{15}A\underline{v}_2 = \dfrac{1}{5}(4,3)$. So

$$\begin{pmatrix} 14 & 16 & 4 \\ -2 & -13 & -22 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} 30 & 0 & 0 \\ 0 & 15 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

We now can write $A = 30\underline{u}_1\underline{v}_1^\top + 15\underline{u}_2\underline{v}_2^\top = \begin{pmatrix} 6 & 12 & 12 \\ -8 & -16 & -16 \end{pmatrix} + \begin{pmatrix} 8 & 4 & -8 \\ 6 & 3 & -6 \end{pmatrix}$.

3.2.11 We finish this section with an example of a 2592×1944 matrix in the form of a picture, and a rank 45 approximation using the SVD:

Theoretically the cost of storing the second picture should be about 4% of the original one, in practice the detailed picture is 836 kB, the fuzzy one 327 kB, so don't always believe what salesmen tell you.

3.3 The Jordan normal form

3.3.1 Associated to a linear map $\mathcal{A}$ of an n-dimensional vector space V to itself is a matrix $A = A_\alpha$ once we have fixed a basis α. If $B = A_\beta$ is the matrix w.r.t. a second basis β, then A and B are related by $B = S^{-1}AS$. Here $S = {}_\alpha S_\beta$ is the transition matrix belonging to the change of β– to α–coordinates. When two matrices are related like this we also call the matrices A and B conjugates of each other (this has nothing to do with complex conjugation!). We leave it as an exercise to the reader that being conjugates is an equivalence relation.

3.3.2 A matrix A is diagonalisable if it is conjugated with a diagonal matrix, and two diagonalisable matrices are conjugated if and only if they are conjugated with the same diagonal matrix. Also, two diagonal matrices are conjugated if and only if they have the same multiset of diagonal elements, so the only difference is the order in which the elements occur.

Not all matrices are diagonalisable however, in this sectie we will look at a normal form that applies to all square matrices, that tells us precisely when two matrices are conjugated. In case of a diagonalisable matrix, this normal form is just a corresponding diagonal matrix. Building blocks of this normal form are the so called Jordan blocks. The field we work in is that of the complex numbers.

3.3.3 Definition. (Jordan block) A Jordan block $B = J_n(\mu)$ is an $n \times n$ matrix with μ's on the diagonal, 1's directly above the diagonal, and 0's elsewhere, so $B_{ij} = \mu$ if $j = i$, 1 if $j = i + 1$ and 0 otherwise.

So, for example $J_3(2) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. The matrix $B = J_n(\mu)$ has the following extremal property: It characteristic polynomial is $(\lambda - \mu)^n$, so B has only one eigenvalue, with algebraic multiplicity n, the geometric multiplicity however, the dimension of the eigenspace E_μ is only 1.

3.3.4 Definition. (Jordan normal form) We say that an $n \times n$ matrix J is in Jordan normal form, if it consists of a sequence of Jordan blocks along the diagonal, and zero's elsewhere.

So J looks as follows:

$$J = \begin{pmatrix} B_1 & O & \\ & \ddots & \\ O & & B_m \end{pmatrix}$$

where, for some n_i and λ_i:

$$B_i = J_{n_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

3.3.5 Theorem. *Every $n \times n$ complex matrix A (so $A \in M_n(\mathbb{C})$) can be brought in Jordan normal form (Jordan form) by conjugation, in other words, there exists an invertible matrix S such that*

$$S^{-1}AS = J = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_m \end{pmatrix}$$

where B_i is a Jordan block of size n_i for eigenvalue λ_i. In particular $n = \sum_{i=1}^m n_i$.

3.3.6 For a diagonalisable matrix all Jordan blocks have size 1. There can be several Jordan blocks for the same eigenvalue. The number of Jordan blocks for the eigenvalue λ equals the dimension of the corresponding eigenspace E_λ, the null space $\mathcal{N}(A-\lambda I)$. We also see that $\dim \mathcal{N}(A-\lambda I)^2 - \dim \mathcal{N}(A-\lambda I)$ equals the the number of Jordan blocks for λ of size at least 2, and so on: $\dim \mathcal{N}(A-\lambda I)^k - \dim \mathcal{N}(A-\lambda I)^{k-1}$ is the number of Jordan blocks (for λ) of size at least k. This implies that the Jordan form of the matrix is uniquely determined (more precisely, the multiset of Jordan blocks).

The Jordan form gives us immediately the characteristic polynomial of the matrix A: $\det(\lambda I - A) = \det(\lambda I - J) = \prod_{i=1}^m (\lambda - \lambda_i)^{n_i}$. But not conversely, the characteristic polynomial does not determine the Jordan form, unless all eigenvalues are different (and then the matrix is diagonalisable. The simplest case is given by the matrices $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, both having λ^2 as characteristic polynomial. To determine the Jordan form we only have to determine the dimensions of the null spaces $N_k = \mathcal{N}(A-\lambda I)^k$. How to find a corresponding basis leads us too far.

3.4 Exercises

§1

1 Determine a LU-decomposition of the following matrices:

a. $A = \begin{pmatrix} 2 & 1 & 3 \\ 6 & 11 & 21 \\ 12 & 26 & 68 \end{pmatrix}$.

b. $B = \begin{pmatrix} 14 & 7 & 0 \\ -6 & -2 & 1 \\ -2 & 0 & 5 \end{pmatrix}$.

2 Show that the following matrices have no (direct) LU-decomposition. Give a permutation of the rows such that the resulting matrix does have a LU-decomposition.

a. $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 5 \end{pmatrix}$.

b. $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 4 \end{pmatrix}$.

3 Let A be the non-singular matrix $= \begin{pmatrix} a & b & 0 \\ c & d & e \\ 0 & 1 & g \end{pmatrix}$. Show that A has a (direct) LU-decomposition if and only if $a \neq 0$ *and* $ad \neq bc$. Show that in that case there is a unique decomposition in which the diagonal of L consists of ones.

4 This was an attempt to make a challanging exercise for LU-decompositions, it kind of failed, so go ahead, see why it is wrong and try to fix it, or ignore it. If the matrix L only has non-zero elements on or directly beneath the diagonal, and the matrix U only has non-zero elements on or directly above the diagonal, the the product $A = LU$ has the property that all non-zero elements are on, directly above or directly beneath the diagonal, in other words $a_{ij} = 0$ if $|i-j| > 1$. Is also the converse true, more precisely, if a *non-singular* matrix A has the property that $a_{ij} = 0$ whenever $|i-j| > 1$, and $A = LU$ is an LU-decomposition of A, then $l_{ij} = 0$ for $i - j > 1$ and $u_{ij} = 0$ for $j - i > 1$?

§2

5 Determine the SVD of the following matrices and write them in the form $\sum \sigma_j \underline{u}_j \underline{v}_j^\top$:

a. $A = \begin{pmatrix} 4 & 0 \\ 0 & 0 \\ 0 & -3 \end{pmatrix}$.

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

c. $A = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 0 & 1 & 2 & 3 \end{pmatrix}$.

6 Use the mathematica command SingularValueDecomposition to find a good rank 2 approximation of the following matrix: $\begin{pmatrix} 0 & 2 & 3 & 4 \\ 2 & 2 & 4 & 5 \\ 3 & 4 & 4 & 6 \\ 4 & 5 & 6 & 6 \end{pmatrix}$.
Hint: put a dot behind one of the numbers, so that mathematica doesn't try to give exact answers.

§3

7 On a 6-dimensional vector space V with basis $\alpha = \{\underline{a}_i : 1 \leq i \leq 6\}$ the map $\mathcal{A}$ is defined by $\mathcal{A}\underline{a}_{2j-1} = (j-1)\underline{a}_{2j-1}$ and $\mathcal{A}\underline{a}_{2j} = (j-1)\underline{a}_{2j} + \underline{a}_{2j-1}$ $(1 \leq j \leq 3)$. Show that the matrix A_α is in Jordan normal form, determine the characteristic polynomial of $\mathcal{A}$, the eigenvalues, and the dimensions of the eigenspaces.

8 Determine the Jordan normal form of the following matrices:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

9 Matrices A and B are conjugate if there is a non-singular matrix S such that $A = S^{-1}BS$, or, what is the same: if they have the same Jordan normal form.

a. Are the following two matrices conjugate?

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ 9 & 7 & 2 \end{pmatrix},$$

b. Same question for the following two matrices

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

Chapter 4

Differential equations

4.1 Systems of differential equations

4.1.1 In this section we apply our knowledge of eigenvalues and eigenvectors to solve systems of differential equations. The functions that appear in these systems are functions of a real variable t. The derivative will be denoted by either a dot above the function symbol, or with the usual prime, so for example $x = x(t)$ and then $\dot{x}$ or x'.

The simplest case is that of a system of one equation in one unknown. Such a system looks as follows

$$\dot{x} = ax + f \ ,$$

here x and f are complex valued functions of a real variabele t and a is a (complex) constant. One way to solve this system is the method of the integrating factor and we find

$$x(t) = ce^{at} + p(t), \quad c \in \mathbb{C},$$

here $\langle e^{at} \rangle$ is the set of all solutions of the corresponding homogeneous equation $\dot{x} = ax$, and $p(t)$ a particuliar solution. If the function $f(t)$ is of the form simple exponential times a polynomial, and this will almost always be the case, then the function $p(t)$ will be the same simple exponential times a (usually different) polynomial.

We first discuss a very simple system of three first order equations in three unknown functions $x(t), y(t)$ and $z(t)$.

4.1.2 Example. Each of the following three equations

$$\begin{aligned}\dot{x} &= 2x + e^t,\\ \dot{y} &= -y,\\ \dot{z} &= 3z + te^{3t}.\end{aligned}$$

can be solved independently, without using the other two equations. Such a system is also called a *decoupled system*. The solution is

$$x = c_1e^{2t} - e^t, \quad y = c_2e^{-t}, \quad z = c_3e^{3t} + \tfrac{1}{2}t^2e^{3t}.$$

We also may write this solution in vector form: $\underline{x} =$

$$c_1\begin{pmatrix}1\\0\\0\end{pmatrix}e^{2t} + c_2\begin{pmatrix}0\\1\\0\end{pmatrix}e^{-t} + c_3\begin{pmatrix}0\\0\\1\end{pmatrix}e^{3t} - \begin{pmatrix}1\\0\\0\end{pmatrix}e^t + \tfrac{1}{2}\begin{pmatrix}0\\0\\1\end{pmatrix}t^2e^{3t}.$$

As a rule we will find the solutions of a system of differential equations in this form.

4.1.3 Example. Consider the system

$$\begin{aligned}\dot{x} &= x + 2y + e^t,\\ \dot{y} &= 12x - y + 2e^t.\end{aligned}$$

When we write this system in matrix form we get:

$$\begin{pmatrix}x\\y\end{pmatrix}' = \begin{pmatrix}1 & 2\\12 & -1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} + \begin{pmatrix}1\\2\end{pmatrix}e^t\ .$$

We try to bring the coefficient matrix A in diagonal form. To do this we need a basis of eigenvectors. The characteristic equation is

$$\begin{vmatrix}1-\lambda & 2\\12 & -1-\lambda\end{vmatrix} = \lambda^2 - 25 = 0\ .$$

So we have two real eigenvalues, 5 and -5. We find eigenspaces $E_5 = \langle(1,2)\rangle$ and $E_{-5} = \langle(-1,3)\rangle$. Take the basis $\alpha = \{(1,2),\ (-1,3)\})$. Then

$${}_\varepsilon S_\alpha = \begin{pmatrix}1 & -1\\2 & 3\end{pmatrix}, \quad {}_\alpha S_\varepsilon = {}_\varepsilon S_\alpha^{-1} = \tfrac{1}{5}\begin{pmatrix}3 & 1\\-2 & 1\end{pmatrix},$$

and we get the diagonal matrix

$$A_\alpha = \begin{pmatrix}5 & 0\\0 & -5\end{pmatrix}.$$

We now transform coordinates to the basis α. So let

$$\begin{pmatrix} u \\ v \end{pmatrix} = {}_{\alpha}S_{\varepsilon} \begin{pmatrix} x \\ y \end{pmatrix} = \tfrac{1}{5} \begin{pmatrix} 3x + y \\ -2x + y \end{pmatrix}.$$

Then $\begin{pmatrix} x \\ y \end{pmatrix} = {}_{\varepsilon}S_{\alpha} \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}' = {}_{\varepsilon}S_{\alpha} \begin{pmatrix} u \\ v \end{pmatrix}'$ and het system changes into

$${}_{\varepsilon}S_{\alpha} \begin{pmatrix} u \\ v \end{pmatrix}' = A_{\varepsilon}S_{\alpha} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^t .$$

Multiplying on the left with ${}_{\alpha}S_{\varepsilon}$ yields

$$\begin{aligned} \begin{pmatrix} u \\ v \end{pmatrix}' &= {}_{\alpha}S_{\varepsilon}A_{\varepsilon}S_{\alpha} \begin{pmatrix} u \\ v \end{pmatrix} + {}_{\alpha}S_{\varepsilon} \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^t, \\ \begin{pmatrix} u \\ v \end{pmatrix}' &= \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t. \end{aligned}$$

This is an decoupled system:

$$u' = 5u + e^t, v' = -5v,$$

and we find the solution straight away:

$$\begin{pmatrix} u \\ v \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{5t} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-5t} - \tfrac{1}{4} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t.$$

Multiplying on the left with ${}_{\varepsilon}S_{\alpha}$ now yields

$${}_{\varepsilon}S_{\alpha} \begin{pmatrix} u \\ v \end{pmatrix} = c_{1\varepsilon}S_{\alpha} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{5t} + c_{2\varepsilon}S_{\alpha} \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-5t} - \tfrac{1}{4}{}_{\varepsilon}S_{\alpha} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t,$$

and we get

$$\begin{pmatrix} x \\ y \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{5t} + c_2 \begin{pmatrix} -1 \\ 3 \end{pmatrix} e^{-5t} - \tfrac{1}{4} \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^t.$$

Notice that the "vector coefficients" of e^{5t} and e^{-5t} are exactly the eigenvectors for 5 and -5.

4.1.4 Definition. A system of n linear differential equations with constant coefficients in n unknown functions $x_1, \ldots, x_n$ of a real variabele t has the form

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}' = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix},$$

here A is an $n \times n$ matrix and $f_1, \ldots, f_n$ are functions of t. This system is called *homogeneous* if $(f_1, \ldots, f_n) = (0, \ldots, 0)$ and *inhomogeneous* otherwise.

4.1.5 Solving such a system boils down to solving a vector equation $\mathcal{B}\underline{x} = \underline{f}$ where $\mathcal{B} : V \to V$ is a linear map. Let the set V consist of sequences of n functions $x_1, \ldots, x_n : \mathbb{R} \to \mathbb{C}$. For convenience we assume that these functions can be differentiated as often as we want.
It is easy to check that V is a (complex) vector space. For every vector $\underline{x} = (x_1, \ldots, x_n) \in V$ we define

$$\mathcal{A}\underline{x} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ and } \mathcal{D}\underline{x} = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} .$$

$\mathcal{A} : V \to V$ and $\mathcal{D} : V \to V$ are linear maps, and the system of differential equations can be written as

$$(\mathcal{D} - \mathcal{A}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} .$$

All solutions of the system are now found by adding to one particular solution all vectors of the null space, so all solutions of the corresponderending homogeneous equation.

Solving a system of linear differential equations typically requires three steps: We first determine all solutions of the homogeneous system, then we find one particular solution and finally we add them up.

4.1.6 We start by discussing how to solve a homogeneous system

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} , \text{ or in vectorform } \underline{x}' = \mathcal{A}\underline{x} .$$

In case we can bring the matrix A in diagonal form, we can decouple the system exactly as in example 4.1.3. We solve the diagonal system, and transform back. This way we find all solutions:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = c_1 \begin{pmatrix} v_{11} \\ \vdots \\ v_{1n} \end{pmatrix} e^{\lambda_1 t} + c_2 \begin{pmatrix} v_{21} \\ \vdots \\ v_{2n} \end{pmatrix} e^{\lambda_2 t} + \cdots + c_n \begin{pmatrix} v_{n1} \\ \vdots \\ v_{nn} \end{pmatrix} e^{\lambda_n t} ,$$

in other words

$$\underline{x} = c_1\underline{v}_1 e^{\lambda_1 t} + c_2\underline{v}_2 e^{\lambda_2 t} + \cdots + c_n\underline{v}_n e^{\lambda_n t} \ ,$$

where $\{\underline{v}_1, \ldots, \underline{v}_n\}$ is a basis of eigenvectors of A (in $\mathbb{C}^n$) with eigenvalues $\lambda_1, \ldots, \lambda_n$.

4.1.7 Example. Consider the homogeneous system

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} -5 & 5 & 7 \\ -2 & 2 & 3 \\ -4 & 5 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \ .$$

The matrix has eigenvalues $1, i$ and $-i$ and we find $E_1 = \langle (2,1,1) \rangle$, $E_i = \langle (10-5i, 4-3i, 5) \rangle$ and $E_{-i} = \langle (10+5i, 4+3i, 5) \rangle$. All solutions are therefore

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = c_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} e^t + c_2 \begin{pmatrix} 10-5i \\ 4-3i \\ 5 \end{pmatrix} e^{it} + c_3 \begin{pmatrix} 10+5i \\ 4+3i \\ 5 \end{pmatrix} e^{-it}$$

with $c_1, c_2, c_3 \in \mathbb{C}$. In order to obtain all real solutions we have to rewrite the last two terms. Write $\underline{a} = (10-5i, 4-3i, 5)e^{it}$ and $\underline{b} = (10+5i, 4+3i, 5)e^{-it}$. Then $\underline{b}$ is the complex conjugate of $\underline{a}$. It follows from this that the span

$$\langle \underline{a}, \underline{b} \rangle = \langle \tfrac{1}{2}(\underline{a}+\underline{b}), \tfrac{1}{2i}(\underline{a}-\underline{b}) \rangle = \langle \mathrm{Re}(\underline{a}), \mathrm{Im}(\underline{a}) \rangle \ .$$

So the set of all solutions of our system can also be described as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = c_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} e^t + c_2 \begin{pmatrix} 10\cos t + 5\sin t \\ 4\cos t + 3\sin t \\ 5\cos t \end{pmatrix} + c_3 \begin{pmatrix} -5\cos t + 10\sin t \\ -3\cos t + 4\sin t \\ 5\sin t \end{pmatrix}$$

with $c_1, c_2, c_3 \in \mathbb{C}$. The real solutions we get by taking $c_1, c_2, c_3 \in \mathbb{R}$.

4.1.8 Consider a homogeneous linear system, in vector notation

$$\underline{x}' = \mathcal{A}\underline{x} \ .$$

The solution method above works if there is a basis of eigenvectors, and in a complex vector space this is usually the case. We now investigate what can be done if no such basis exists. We will limit ourselves to the simplest case: the characteristic equation has a root α with multiplicity 2, but the corresponding eigenspace only has dimension 1.

In this case we look for solutions of the form

$$\underline{x} = (\underline{u}t + \underline{v})e^{\alpha t} \ .$$

Substituting this into the equation gives

$$(\alpha\underline{u}t + \alpha\underline{v} + \underline{u})e^{\alpha t} = (\mathcal{A}\underline{u}\ t + \mathcal{A}\underline{v})e^{\alpha t} \ ,$$

and from this we get

$$\alpha\underline{u} = \mathcal{A}\underline{u} \ \text{ and } \ \alpha\underline{v} + \underline{u} = \mathcal{A}\underline{v} \ .$$

One solution we see immediately: take $\underline{u} = \underline{0}$ and $\underline{v}$ an eigenvector with eigenvalue α.

But, there is also a solution with $\underline{u} \neq \underline{0}$. This is a consequence of 1.5.14: since the the dimension of the eigenspace is smaller than the multiplicity of the eigenvalue there exist vectors $\underline{u} \neq \underline{0}, \underline{v} \neq \underline{0}$ such that $(\mathcal{A} - \alpha I)\underline{v} = \underline{u}(\neq \underline{0})$ and $(\mathcal{A} - \alpha I)\underline{u} = \underline{0}$, that is $\mathcal{A}\underline{u} = \alpha\underline{u}$ and $\mathcal{A}\underline{v} = \alpha\underline{v} + \underline{u}$.

4.1.9 Example. Consider the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 5 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} .$$

The characteristic equation is $(\lambda - 3)^2 = 0$, so $\lambda = 3$ is eigenvalue with multiplicity 2, but we find $E_3 = \langle (1,1) \rangle$. We now look for solutions of the form $\underline{x} = (\underline{u}\, t + \underline{v})e^{3t}$. Substitution in the equation gives

$$(3\underline{u}t + 3\underline{v} + \underline{u})e^{3t} = (\mathcal{A}\underline{u}\ t + \mathcal{A}\underline{v})e^{3t} \ ,$$

so $\mathcal{A}\underline{u} = 3\underline{u}$ and $\mathcal{A}\underline{v} = 3\underline{v} + \underline{u}$. If $\underline{u} = \underline{0}$, then $\mathcal{A}\underline{v} = 3\underline{v}$, so we take $\underline{v} = (1,1)$; one solution is therefore (1,1) e^{3t}. $\mathcal{A}\underline{u} = 3\underline{u}$ is satisfied by $\underline{u} = (1,1)$. Now solve $\underline{v}$ from $(\mathcal{A} - 3\mathcal{I})\underline{v} = \underline{u}$:

$$\left(\begin{array}{cc|c} 2 & -2 & 1 \\ 2 & -2 & 1 \end{array}\right) .$$

This system has solutions: $2v_1 - 2v_2 = 1$, so take for instance $\underline{v} = (\frac{1}{2}, 0)$. A second solution to the homogeneous system of differential equations is therefore $\underline{x} = ((1,1)\, t + (\frac{1}{2}, 0))e^{3t}$. The solutions we find are now

$$\begin{pmatrix} x \\ y \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{3t} + c_2 \begin{pmatrix} t + \frac{1}{2} \\ t \end{pmatrix} e^{3t} .$$

4.1.10 We look at the problem of finding a particular solution for an inhomogeneous system of linear differential equations.

For this we use, if appropriate, the so called *superposition principle* : If $\underline{p}_1$ is a particular solution of $\underline{x}' = \mathcal{A}\underline{x} + \underline{f}_1$ and if $\underline{p}_2$ is a particular solution of $\underline{x}' = \mathcal{A}\underline{x} + \underline{f}_2$, then $\underline{p}_1 + \underline{p}_2$ is a particular solution of $\underline{x}' = \mathcal{A}\underline{x} + (\underline{f}_1 + \underline{f}_2)$. This property follows directly from the linearity of differentiation and of $\mathcal{A}$.

As a theoretical introduction we consider system

$$\underline{x}' = \mathcal{A}\underline{x} + \underline{a}e^{\alpha t}$$

for some constant vector $\underline{a}$. In this situation we try to find a particular solution of the form $\underline{x} = \underline{u}e^{\alpha t}$. Substition in the equation leads to

$$\alpha \underline{u} e^{\alpha t} = \mathcal{A}\underline{u}\ e^{\alpha t} + \underline{a}e^{\alpha t}\ ,$$

$$(\mathcal{A} - \alpha\mathcal{I})\underline{u} = -a\ .$$

So we will find such a vector $\underline{u}$ if $-\underline{a}$ (and therefore $\underline{a}$) is contained in the range of $\mathcal{A} - \alpha\mathcal{I}$.

If α is not an eigenvalue, then the map $\mathcal{A} - \alpha\mathcal{I}$ is invertible and the equation has a unique solution for every a. If α *is* an eigenvalue, then either there are infinitely many solutions for $\underline{u}$ (namely in the case that $-\underline{a}$ happens to be contained in the range of $\mathcal{A} - \alpha\mathcal{I}$) or no solution, if this is not the case.

In the second case we try to find a particular solution of the type $\underline{x} = (\underline{u}t + \underline{v})e^{\alpha t}$. When we substitute this in the equation we get

$$(\underline{u} + \alpha\underline{v} + \alpha\underline{u}t)e^{\alpha t} = (t\mathcal{A}\underline{u} + \mathcal{A}\underline{v} + \underline{a})e^{\alpha t}\ ,$$

$$\mathcal{A}\underline{u} = \alpha\underline{u}, \quad \mathcal{A}\underline{v} + \underline{a} = \underline{u} + \alpha\underline{v}\ ,$$

$$(\mathcal{A} - \alpha\mathcal{I})\underline{u} = \underline{0}, \quad (\mathcal{A} - \alpha\mathcal{I})\underline{v} = \underline{u} - \underline{a}\ .$$

Our job is now to find an eigenvector $\underline{u}$ for which $\underline{u} - \underline{a}$ is contained in the range of $\mathcal{A} - \alpha\mathcal{I}$. Usually this is possible: Let $\mathcal{R}$ and $\mathcal{N}$ be the range and the null space of $\mathcal{A} - \alpha\mathcal{I}$. If $\mathcal{R} \cap \mathcal{N} = \{\underline{0}\}$, then (since $\dim \mathcal{R} + \dim \mathcal{N} = n$) every vector $\underline{a}$ can be written as $\underline{a} = \underline{u} + \underline{w}$ with $\underline{u} \in \mathcal{N}$ and $\underline{w} \in \mathcal{R}$. In this case $\underline{u} - \underline{a} = -\underline{w} \in \mathcal{R}$, so that $\underline{v}$ exists. Notice that the condition $\mathcal{R} \cap \mathcal{N} = \{\underline{0}\}$ means that the multiplicity of α as solution of the characteristic equation is equal to the dimension of the corresponding eigenspace. In the unlikeley case that we still not find a particular solution, we try $\underline{x} = (\underline{u}t^2 + \underline{v}t + \underline{w})e^{\alpha t}$, etc. It can be shown that we will find a solution in this way eventually.

To find a particular solution of the system

$$\underline{x}' = \mathcal{A}\underline{x} + \underline{p}(t)e^{\alpha t}$$

with $\underline{p}(t) = \underline{a}_0 + \underline{a}_1 t + \cdots + \underline{a}_n t^n$ we try $\underline{x} = \underline{q}(t)e^{\alpha t}$ with $\underline{q}(t) = \underline{b}_0 + \underline{b}_1 t + \cdots + \underline{b}_n t^n$. If this does not work, we try again, this time with a (vector) polynomial of degree $n+1$, etc.

4.1.11 Example. We look for a particular solution of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}' = \begin{pmatrix} 0 & -8 & 8 \\ 1 & 4 & -1 \\ 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 6 \\ -2 \\ -4 \end{pmatrix} e^{2t} .$$

Let A be the coefficient matrix and $\underline{a} = (6, -2, -4)$.

We look for a particular solution of the form $\underline{x} = \underline{u}e^{2t}$. Substition of this in the equation gives

$$2\underline{u}e^{2t} = \mathcal{A}\underline{u}e^{2t} + \underline{a}e^{2t} ,$$

$$(\mathcal{A} - 2\mathcal{I})\underline{u} = -\underline{a} ,$$

$$\left(\begin{array}{ccc|c} -2 & -8 & 8 & -6 \\ 1 & 2 & -1 & 2 \\ 0 & 2 & -4 & 4 \end{array}\right) ,$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & -3 \end{array}\right) .$$

So $\underline{u} = (7, -4, -3)$ and a particular solution is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ -4 \\ -3 \end{pmatrix} e^{2t} .$$

4.1.12 Example. We look for a particular solution of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}' = \begin{pmatrix} 0 & -8 & 8 \\ 1 & 4 & -1 \\ 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \left(t \begin{pmatrix} -22 \\ 3 \\ 10 \end{pmatrix} + \begin{pmatrix} 21 \\ -6 \\ -4 \end{pmatrix} \right) e^{2t} .$$

Let A be the coefficient matrix, $\underline{a} = (-22, 3, 10)$, $\underline{b} = (21, -6, -4)$. We try a solution of the form $\underline{x} = (\underline{u}t + \underline{v})e^{2t}$. Substitution in the equation gives

$$(\underline{u} + 2\underline{u}t + 2\underline{v})e^{2t} = (\mathcal{A}\underline{u}\ t + \mathcal{A}\underline{v} + \underline{a}t + \underline{b})e^{2t} ,$$

$$\mathcal{A}\underline{u} + \underline{a} = 2\underline{u},\ \underline{u} + 2\underline{v} = \mathcal{A}\underline{v} + \underline{b} ,$$

$$(\mathcal{A} - 2\mathcal{I})\underline{u} = -\underline{a},\ (\mathcal{A} - 2\mathcal{I})\underline{v} = \underline{u} - \underline{b} .$$

We start by solving the equation for $\underline{u}$:

$$\left(\begin{array}{rrr|r} -2 & -8 & 8 & 22 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -4 & -10 \end{array}\right) \quad \to \quad \left(\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right),$$

we find $\underline{u} = (1, -1, 2)$. Next, with this $\underline{u}$ we solve the equation $\underline{v}$:

$$\left(\begin{array}{rrr|r} -2 & -8 & 8 & -20 \\ 1 & 2 & -1 & 5 \\ 0 & 2 & -4 & 6 \end{array}\right) \quad \to \quad \left(\begin{array}{rrr|r} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right).$$

So $\underline{v} = (2, 1, -1)$ and a particular solution is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \left(t \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \right) e^{2t} \; .$$

4.1.13 Example. Consider the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 1 & -2 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \sin t \\ \cos t \end{pmatrix} \; .$$

We first solve the homogeneous system. The characteristic equation is $\lambda^2 + 9 = 0$, so the eigenvalues are $3i$ and $-3i$. We find the eigenspaces $E_{3i} = \langle (1+3i, 5) \rangle$ and $E_{-3i} = \langle (1-3i, 5) \rangle$. All solutions of the homogeneous equation are therefore

$$\begin{pmatrix} x \\ y \end{pmatrix} = c_1 \begin{pmatrix} 1+3i \\ 5 \end{pmatrix} e^{3it} + c_2 \begin{pmatrix} 1-3i \\ 5 \end{pmatrix} e^{-3it} \; ,$$

or, in real form:

$$\begin{pmatrix} x \\ y \end{pmatrix} = c_1' \begin{pmatrix} \cos 3t - 3 \sin 3t \\ 5 \cos 3t \end{pmatrix} + c_2' \begin{pmatrix} 3 \cos 3t + \sin 3t \\ 5 \sin 3t \end{pmatrix} \; .$$

In order to find a particular solution we start by noticing that

$$\begin{pmatrix} \sin t \\ \cos t \end{pmatrix} = \mathrm{Im} \begin{pmatrix} 1 \\ i \end{pmatrix} e^{it} \; .$$

We now look for a particular solution of

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 1 & -2 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ i \end{pmatrix} e^{it} \; .$$

Since the matrix is real and since $(\mathrm{Im}\ \underline{x})' = \mathrm{Im}(\underline{x}')$, the imaginary part of this particular solution is a particular solution of the original system.

We try therefore $\underline{x} = \underline{u}e^{it}$. Substituting in the system gives $\underline{u}ie^{it} = (\mathcal{A}\underline{u} + \underline{a})e^{it}$, in other words $(\mathcal{A} - i\mathcal{I})\underline{u} = -\underline{a}$. From the system

$$\left(\begin{array}{cc|c} 1-i & -2 & -1 \\ 5 & -1-i & -i \end{array}\right)$$

now follows $\underline{u} = \frac{1}{8}(1-i, 4-i)$.

A particular solution of the second system is therefore

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 1-i \\ 4-i \end{pmatrix} e^{it}$$

with imaginary part

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{8}\begin{pmatrix} -\cos t + \sin t \\ -\cos t + 4\sin t \end{pmatrix}.$$

Hence all real solutions of the system are given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{8}\begin{pmatrix} -\cos t + \sin t \\ -\cos t + 4\sin t \end{pmatrix} +$$

$$c_1 \begin{pmatrix} \cos 3t - 3\sin 3t \\ 5\cos 3t \end{pmatrix} + c_2 \begin{pmatrix} 3\cos 3t + \sin 3t \\ 5\sin 3t \end{pmatrix}$$

with c_1 and $c_2 \in \mathbb{R}$.

4.1.14 We close this section with the remark that in practice we often not only have the system of differential equations but also several (boundary) conditions that have to be satisfied by the solutions. In such cases we first determine the general solution of the system, en then we try to determine the constants that occur in such a way that the conditions are satisfied.

4.2 Notes

Systems of differential equations describe situations where several quantities influence one and other. In the context of systems of differential equations, eigenvalues and eigenvectors were introduced to decouple such a system.

4.3 Exercises

§1

1 Solve the following systems of linear differential equations

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} f_1 \\ f_2 \end{pmatrix},$$

where $\underline{x} = \underline{x}(t)$ in some cases has to satisfy an extra condition.

a. $\underline{x}' = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix} \underline{x}$,

b. $\underline{x}' = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix} \underline{x}$, $\underline{x}(0) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$,

c. $\underline{x}' = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix} \underline{x} + \begin{pmatrix} 8 \\ 7 \end{pmatrix} e^t$,

d. $\underline{x}' = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix} \underline{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-t}$,

e. $\underline{x}' = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix} \underline{x} + \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$.

2 Solve the following systems of linear differential equations

$$\underline{x}' = A\underline{x} + \underline{f},$$

where $\underline{x} = \underline{x}(t)$ in some cases has to satisfy an extra condition.

a. $\underline{x}' = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \underline{x} + \begin{pmatrix} 1 \\ 5 \end{pmatrix}$,

b. $\underline{x}' = \begin{pmatrix} 1 & -2 \\ 5 & 3 \end{pmatrix} \underline{x}$,

c. $\underline{x}' = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \underline{x}$,

d. $\underline{x}' = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \underline{x} + \begin{pmatrix} 4 \\ 1 \end{pmatrix} e^t$, $\underline{x}(0) = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$,

e. $\underline{x}' = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \underline{x}$, $\lim_{t\to\infty} \underline{x}(t) = \underline{0}$,

f. $\underline{x}' = \begin{pmatrix} 2 & -4 \\ 1 & 6 \end{pmatrix} \underline{x}$,

g. $\underline{x}' = \begin{pmatrix} 2 & -4 \\ 1 & 6 \end{pmatrix} \underline{x} + \begin{pmatrix} 2 \\ -7 \end{pmatrix}$, $\underline{x}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

3 Determine the solutions of the following systems of linear differential equations.

a. $\begin{pmatrix} 1 & 1 \\ 8 & 3 \end{pmatrix} \underline{x}' = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \underline{x}$,

b. $\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \underline{x}' = \begin{pmatrix} 1 & 0 \\ 3 & 5 \end{pmatrix} \underline{x}$,

c. $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \underline{x}' = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \underline{x} + \begin{pmatrix} 5 \\ 4 \end{pmatrix} e^t$.

4 Solve the following systems of linear differential equations

$$\underline{x}' = A\underline{x} + \underline{f},$$

where $\underline{x} = \underline{x}(t)$ in some cases has to satisfy an extra condition.

a. $\underline{x}' = \begin{pmatrix} 1 & -1 & -1 \\ -2 & 2 & -1 \\ 2 & 0 & 3 \end{pmatrix} \underline{x}$,

b. $\underline{x}' = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \underline{x}$,

c. $\underline{x}' = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \underline{x} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} e^{4t}$, $\underline{x}(0) = \frac{1}{3} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$,

d. $\underline{x}' = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \underline{x} + \left(\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} t \right) e^{2t}$,

e. $\underline{x}' = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 2 & 6 & -2 \end{pmatrix} \underline{x}$,

f. $\underline{x}' = \frac{1}{3}\begin{pmatrix} 8 & 0 & -5 \\ -4 & 6 & 1 \\ -1 & -3 & 7 \end{pmatrix}\underline{x} + \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} e^{3t}, \quad \underline{x}(0) = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix},$

g. $\underline{x}' = \frac{1}{3}\begin{pmatrix} 8 & 0 & -5 \\ -4 & 6 & 1 \\ -1 & -3 & 7 \end{pmatrix}\underline{x} + \begin{pmatrix} 4\cos 2t + 3\sin 2t \\ -3\cos 2t - \sin 2t \\ 0 \end{pmatrix}, \quad \underline{x}(0) = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.$

4.3.1 Exercises on exam level

5 Determine the real solution of the system of differential equations

$$\underline{x}' = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}\underline{x} + \begin{pmatrix} 3 \\ 2 \end{pmatrix} e^{t} \quad \text{with} \quad \underline{x}(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

6 Determine all solutions of the system of differential equations

$$\underline{x}' = \begin{pmatrix} 2 & 2 \\ 3 & 1 \end{pmatrix}\underline{x} + \begin{pmatrix} 5 \\ 0 \end{pmatrix} e^{4t}.$$

7 Determine alle solutions of the system of differential equations

$$\underline{x}' = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}\underline{x} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} e^{t}.$$

Appendix A

Prerequisites

A.1 Notation

A.1.1 Sets

Sets consist of elements. We denote the fact that a is an element of the set A by:

$$a \in A.$$

Sets are usually specified in one of the following ways:

- **Listing the elements between curly braces**. For example:

  $$\{1, 2, 3, 5\}, \quad \{1, 2, 3, \ldots\}, \quad \{1, 2, 3, 5, 3\}, \quad \{2, \sqrt{3}, x^2 - 1\}.$$

  The dots in the second example mean that we expect the reader to recognize the pattern and complete it: so 4, 5 etc. also belong to this set. Two sets are equal if they contain the same elements, so the first and the third set are equal.

- **A description using a defining property**. Examples:

  $$\{x \mid x \text{ is an even integer}\}, \quad \{y \mid y \text{ is real and } y < 0\}.$$

  More formally we can write

  $$\{x \in \mathbf{Z} \mid x \text{ even}\}, \quad \{y \in \mathbf{R} \mid y < 0\},$$

  making clear in which set (universe) our elements live.

As a reminder we list some frequently used no(ta)tions.

$\emptyset$	the empty set
$a \notin A$	a is *not* an element of A
$A \subset B$ (sometimes: $B \supset A$)	A is a subset of B
(or $A \subseteq B$)	(or: A is contained in B)
	i.e. if $a \in A$ then $a \in B$
$A \not\subset B$	A is not a subset of B
$A \cap B := \{x \mid x \in A \text{ and } x \in B\}$	the *intersection* of A and B
$A \cup B := \{x \mid x \in A \text{ or } x \in B\}$	the *union* of A and B
$A - B := \{x \mid x \in A \text{ and } x \notin B\}$	
(or: $A \setminus B$)	the *set theoretic difference* of A and B
$A \times B := \{(a, b) \mid a \in A,\, b \in B\}$	the (Cartesian) *product* of A and B
$A_1 \times A_2 \times \cdots \times A_n :=$	the *product* of A_1, $A_2, \ldots, A_n$
$\{(a_1, a_2, \ldots, a_n) \mid a_1 \in A_1,$	
$a_2 \in A_2, \ldots, a_n \in A_n\}$	
$A^n := \{(a_1, a_2, \ldots, a_n) \mid$	special case
$a_1, \ldots, a_n \in A\}$	

A.1.2 Maps

For sets A and B, a *map* from A to B is a rule associating to each element of A exactly one element of B. Notation: $f : A \to B$. The set A is called the *domain* of the map, B the *codomain*. If the elements of B are numbers, then one often uses the more common word *function* in stead of map. Two functions are equal if they have the same domain, the same codomain and the same value for every element of the domain.

Some notations:

$f : A \to B$	map with domain A and codomain B (clearly other letters are also allowed!)
$f(a)$	the element associated to a, the value of f in a: the *image* of a (under f)
$f : a \mapsto b$	f maps a to b
$f(D) := \{f(d) \mid d \in D\}$	the *image* of D, for a subset D of A
$f(A)$	special case: the *image or range* of f
$f^{-1}(E) := \{a \in A \mid f(a) \in E\}$ (or: $f^{\leftarrow}(E)$)	the *(complete) inverse image* of E (E a subset of B)
$f^{-1}(b)$ i.s.o. $f^{-1}(\{b\})$	keeps notation simple
$f : A \to B$ *injective* (f is an *injection*)	for all $a, a' \in A$: $f(a) = f(a') \Rightarrow a = a'$; or: for all $a, a' \in A$: if $a \neq a'$ then $f(a) \neq f(a')$
$f : A \to B$ *surjective* (f is a *surjection*)	for all b there is an $a \in A$ with $f(a) = b$, i.o.w. $f(A) = B$
$f : A \to B$ *bijective* (f is a *bijection*)	f is injective and surjective (so: for all $b \in B$ there is a unique $a \in A$ with $f(a) = b$)

If $f : A \to B$ is a bijection, then for all $b \in B$ there is a unique $a \in A$ with $f(a) = b$. In this case we can define a map from B to A by the rule: $b \mapsto a$ if $f(a) = b$. This map (that only exists if f is a bijection) is called the *inverse* of f and is denoted by f^{-1}. But take care, because the same symbol might also refer to the inverse image. The context should make clear what is meant.

A.2 Trigonometric formulas

Some important identities involving the trigonometric functions:

- $\cos^2(x) + \sin^2(x) = 1$;
- $\sin(x + 2\pi) = \sin(x)$ and $\cos(x + 2\pi) = \cos(x)$;
- $\sin(\pi - x) = \sin(x)$ and $\cos(\pi - x) = -\cos(x)$;
- $\sin(\pi + x) = -\sin(x)$ and $\cos(\pi + x) = -\cos(x)$;
- $\sin(\pi/2 - x) = \cos(x)$ and $\cos(\pi/2 - x) = \sin(x)$;

- $\sin(2x) = 2\sin(x)\cos(x)$ and $\cos(2x) = \cos^2(x) - \sin^2(x)$;
- $\sin(x+y) = \sin(x)\cos(y)+\cos(x)\sin(y)$ and $\cos(x+y) = \cos(x)\cos(y)-\sin(x)\sin(y)$.

A.3 The Greek alphabet

In mathematics we use apart from the usual letters and digits also many letters from the Greek alphabet. Here we present a table with the Greek alphabet, letters that are frequently use in Linear Algebra 1 and 2 are indicated with a *.

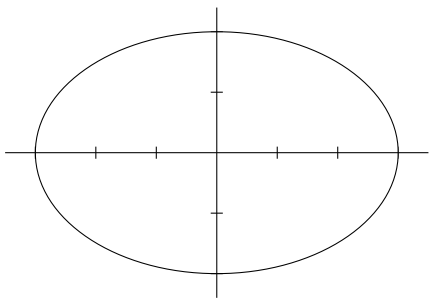
name	**minuscule**	**capital**
alpha	α *	A
beta	β *	B
gamma	γ *	Γ
delta	δ *	Δ
epsilon	ε	E
zeta	ζ	Z
eta	η	H
theta	θ or ϑ	Θ
iota	ι	I
kappa	κ	K
lambda	λ *	Λ
mu	μ *	M
nu	ν	N
xi	ξ	Ξ
omikron	o	O
pi	π	Π
rho	ρ *	R
sigma	σ *	Σ
tau	τ *	T
upsilon	υ	Υ
phi	ϕ or φ *	Φ
chi	χ	X
psi	ψ *	Ψ
omega	ω *	Ω

A.4 Ellipse, hyperbola, parabola

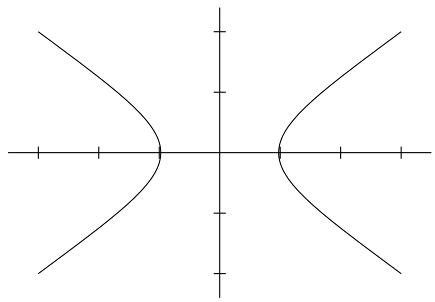
The standard equation for an *ellipse* is $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ where a, b are real (positive) numbers.

The standard equation for a *hyperbola* is $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ with a, b real (positive) numbers.

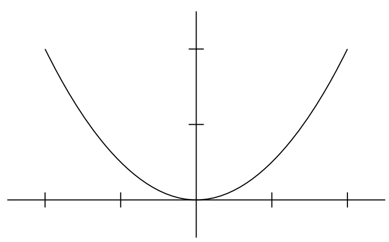
The standard equation for a *parabola* is $y = ax^2$ with a a real number different from 0.



The ellipse $\frac{1}{9}x^2 + \frac{1}{4}y^2 = 1$.



The hyperbola $x^2 - 2y^2 = 1$.



The parabola with equation $y = \frac{1}{2}x^2$.

Appendix B

Recap Linear Algebra 1

B.1 Complex numbers

B.1.1 (fields) A field consist of a set $\mathbb{K}$, with two special elements, 0 and 1, and two binary operations, addition and multiplication $+, \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$. Addition and multiplication, together with 0 and 1 should obey the following rules (we omit most of the $\forall$-quantifiers, and mostly write ab for $a \cdot b$):

(i) $a + b = b + a$ addition is commutative;

(ii) $(a + b) + c = a + (b + c)$ addition is associative;

(iii) $a + 0 = a$ 0 acts as the neutral element for addition;

(iv) $\forall a : \exists(-a) : a + (-a) = 0$ existence of an inverse for addition;

(i)–(iv) say that $(\mathbb{K}, +, 0)$ is an abelian group.

(v) $ab = ba$ multiplication is commutative;

(vi) $(ab)c = a(bc)$ multiplication is associative;

(vii) $1 \cdot a = a$ 1 acts as the neutral element for multiplication;

(viii) $\forall a \neq 0 : \exists a^{-1} : a \cdot a^{-1} = 1$ existence of an inverse for multiplication;

(v)–(viii) say that $(\mathbb{K}^*, \cdot, 1)$, where $\mathbb{K}^* := \mathbb{K} \setminus \{0\}$ is an abelian group;

(ix) $(a + b)c = ac + bc$ multiplication is distributive over addition;

B.1.2 (complex numbers) The fields we know from high school are $\mathbb{Q}$, the rationals, and $\mathbb{R}$ the real numbers. We turn the vector space $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ into a field, $\mathbb{C}$, by means of the following addition and multiplication.
$(a, b) + (c, d) = (a + c, b + d)$, the usual addition in $\mathbb{R}^2$.
$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$; together with $0 = (0, 0)$ and $1 = (1, 0)$ we

may verify that we get a field. Obviously $-(a,b) = (-a,-b)$, less obvious: $(a,b)^{-1} = (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$. Let $i := (0,1)$, since $1 = (1,0)$ we may now write $(a,b) = a+bi$, together with $i^2 = -1$ this gives us back all our rules: $(a+bi)+(c+di) = (a+c)+(b+d)i$ and $(a+bi)(c+di) = ac+adi+bic+bidi = ac-bd+(ad+bc)i$.

B.1.3 Let $z = x+yi$, $x, y \in \mathbb{R}$ be a complex number.

$\mathrm{Re}\ z = x$, the real part of z;

$\mathrm{Im}\ z = y$, the imaginary part of z;

$\bar{z} = x - yi$, the complex conjugate of z;

$|z| = \|(x,y)\| = \sqrt{x^2+y^2}$, the modulus, length, or absolute value of z;

$\bar{0} = 0$; $\bar{1} = 1$; $\overline{z_1+z_2} = \overline{z_1} + \overline{z_2}$; $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$; $z = \bar{z} \Leftrightarrow z \in \mathbb{R}$;

The last line shows that the map $z \mapsto \bar{z}$ is an $\mathbb{R}$-automorphism of $\mathbb{C}$;

We have $z + \bar{z} = 2\ \mathrm{Re}\ z$; $z - \bar{z} = 2i\ \mathrm{Im}\ z$; $z\bar{z} = |z|^2$; $z^{-1} = \bar{z}/|z|^2$;

B.1.4 (polar representation) Every $z \in \mathbb{C}$ has a representation of the form $z = r(\cos\phi + i\sin\phi) =: r\ \mathrm{cis}\ \phi$, $r, \phi \in \mathbb{R}$. Here we take $r = |z| \geq 0$, the modulus of z. If $z = 0$ then ϕ is not determined, if $z \neq 0$ then ϕ represents the angle between the positive x-axis and the vector $(x,y) \in \mathbb{R}^2$, so ϕ is determined upto a multiple of 2π. We call ϕ the *argument* of z, notation $\arg z$. Sometimes we speak of the *principal argument*, by which we mean that ϕ is chosen from a fixed interval of length 2π, usually $[0, 2\pi)$ or $(-\pi, \pi]$. From the addition formulas for cosine and sine we find, that if $z_1 = r_1\ \mathrm{cis}\ \alpha$ and $z_2 = r_2\ \mathrm{cis}\ \beta$ that

$$\begin{aligned} z_1 z_2 &= r_1 r_2\ ((\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\cos\alpha\sin\beta + \sin\alpha\cos\beta)) = \\ &= r_1 r_2\ (\cos(\alpha+\beta) + i\sin(\alpha+\beta)) = r_1 r_2\ \mathrm{cis}(\alpha+\beta). \end{aligned}$$

From this we get the important rules:

$|z_1 z_2| = |z_1||z_2|$, the modulus of the product is the product of the moduli;

$\arg\ z_1 z_2 = \arg\ z_1 + \arg\ z_2$, the argument of the product is the sum of the arguments;

B.1.5 (exponential representation) The function cis behaves very much like the function exp: $\mathrm{cis}\ \alpha \cdot \mathrm{cis}\ \beta = \mathrm{cis}\ (\alpha+\beta)$ and $\exp x \cdot \exp y = \exp(x+y)$. This property allows us to extend the definition of the exponential function to the field $\mathbb{C}$ (and forget about cis). Let $z = x + yi$, $x, y, \alpha \in \mathbb{R}$, then we define:

$$e^{i\alpha} = \text{cis } \alpha, \qquad \exp(z) = e^z = e^{x+yi} = e^x \cdot e^{iy} = e^x \text{cis } y = e^x(\cos y + i \sin y).$$

In other words: $|e^z| = e^x = e^{\text{Re } z}$ and $\arg(e^z) = y = \text{Im } z$ also, if z has modulus r and argument ϕ, then $z = r\, e^{i\phi}$.

B.1.6 (fundamental theorem of algebra) An important reason that complex numbers play a useful role in the theory of real vector spaces is the following theorem.

B.1.7 Theorem. *Every non-constant polynomial over $\mathbb{C}$ has a zero in $\mathbb{C}$.*

This theorem is of fundamental importance, as the name suggests. If $f(z)$ is a polynomial, and $z = a$ is a zero, then $z - a$ divides $f(z)$, that is, $f(z) = (z - a)g(z)$, for some polynomial g. If g is not constant, we may apply the theorem again, and finally we find a factorization of f in linear factors $f(z) = (z - a_1)(z - a_2) \cdots (z - a_n)$ if f has degree n.

B.1.8 (real polynomials) If $p(z) = p_n z^n + p_{n-1} z^{n-1} + \cdots + p_1 z + p_0$ is a (complex) polynomial with real coefficients, so $p_i \in \mathbb{R}$ for all i, then

$$\begin{aligned} \overline{p(z)} &= \overline{p_n z^n + p_{n-1} z^{n-1} + \cdots + p_1 z + p_0} = \\ &= p_n \bar{z}^n + p_{n-1} \bar{z}^{n-1} + \cdots + p_1 \bar{z} + p_0 = p(\bar{z}), \end{aligned}$$

because conjugation is an automorphism of $\mathbb{C}$ that fixes $\mathbb{R}$ elementwise. As a consequence we have that $p(a) = 0$ implies $p(\bar{a}) = \overline{p(a)} = 0$. If a is real this is obvious, since then $\bar{a} = a$, but if $a \neq \bar{a}$ then $p(z)$ is divisible by $(z - a)(z - \bar{a}) = z^2 - (a + \bar{a})z + a\bar{a} = z^2 - 2 \text{ Re } a + |a|^2$ a real, second degree polynomial. So we see that over the reals every polynomial factors in linear and quadratic factors.

B.2 Vectors

B.2.1 (vector space) A vector space consists of a set V, with a special element $\underline{0}$, the zero vector, and two operations, addition: $+ : V \times V \to V$, $(\underline{v}, \underline{w}) \mapsto \underline{v} + \underline{w}$ and scalar multiplication $\cdot : \mathbb{K} \times V \to V$, $(\lambda, \underline{v}) \mapsto \lambda \cdot \underline{v}$ or simply $\lambda \underline{v}$. Addition, scalar multiplication and zero vector should obey the following rules (we omit most of the $\forall$-quantifiers):

(i) $\underline{a} + \underline{b} = \underline{b} + \underline{a}$ addition is commutative;

(ii) $(\underline{a}+\underline{b})+\underline{c}=\underline{a}+(\underline{b}+\underline{c})$ addition is associative;

(iii) $\underline{a}+\underline{0}=\underline{a}$ $\underline{0}$ acts as the neutral element for addition;

(iv) $\forall\underline{a}:\exists(-\underline{a}):\underline{a}+(-\underline{a})=\underline{0}$ existence of an inverse for addition;

(i)–(iv) say that $(V,+,\underline{0})$ is an abelian group.

(v) $\lambda(\underline{a}+\underline{b})=\lambda\underline{a}+\lambda\underline{b}$ multiplication is distributive over addition;

(vi) $(\lambda+\mu)\underline{a}=\lambda\underline{a}+\mu\underline{a}$ another kind of distributivity;

(vii) $(\lambda\mu)\underline{a}=\lambda(\mu\underline{a})$ kind of associativity;

(viii) $1\cdot\underline{a}=\underline{a},\quad(-1)\cdot\underline{a}=-\underline{a},\quad 0\cdot\underline{a}=\underline{0}$ no rules, follows from (i)–(vii).

B.2.2 (the vector space $\mathbb{K}^n$) By $\mathbb{K}^n$ we denote the set of (ordered) n-tuples of elements of $\mathbb{K}$, so $\underline{v}\in\mathbb{K}^n$ means $\underline{v}=(v_1,\ldots,v_n)$. With the following rules for addition and scalar multiplication $\mathbb{K}^n$ is a vector space: $\underline{v}+\underline{w}=(v_1+w_1,\ldots,v_n+w_n)$ and $\lambda\underline{v}=(\lambda v_1,\ldots,\lambda v_n)$. The zero vector in this vector space is $\underline{0}=(0,\ldots,0)$.

B.2.3 (subspace and span) A subset W of V is a (linear) subspace if (i) $\underline{0}\in W$ and (ii) W is closed under addition and scalar multiplication. If $\{\underline{v}_1,\ldots,\underline{v}_m\}$ is a set of vectors in V then the span $<\underline{v}_1,\ldots,\underline{v}_m>$ is the set of all linear combinations $\lambda_1\underline{v}_1+\cdots+\lambda_m\underline{v}_m$. The span of the empty set is by definition $\{\underline{0}\}$. The span of an infinite set is the collection of all *finite* linear combinations, so only a finite number of coefficients λ is non-zero. This finiteness is essential, and actually part of the definition of linear combination. The span of a set A or also the subspace spanned by A is the smallest subspace of V containing all vectors from A.

B.2.4 (linear (in)dependence) A linear combination is called *non-trivial* if at least one of the coefficients is non-zero. A set A of vectors in a vector space V is called *dependent* if there is a non-trivial linear combination that equals the zero vector, otherwise the set is called *independent*. The usual way to prove independence is this: suppose $\lambda_1\underline{v}_1+\cdots+\lambda_m\underline{v}_m=\underline{0}$ then such and so and hence $\lambda_1=\cdots=\lambda_m=0$.

B.2.5 (basis and dimension) If a set of vectors α spans V and is linearly independent, we say that α is a basis for V. Important properties: Every vector space has a basis (if the basis is infinite we need some version of the axiom of choice to prove this). All bases of a vector space have the same size (cardinality), called the *dimension* of V.

B.2.6 (inner product) For $\underline{a}, \underline{b} \in \mathbb{R}^n$ we denote the standard inner product by $(\underline{a}, \underline{b})$, other popular notations are $\underline{a} \bullet \underline{b}$ or $\langle \underline{a}, \underline{b} \rangle$:

$$(\underline{a}, \underline{b}) = a_1 b_1 + \cdots + a_n b_n.$$

We say that $\underline{a}$ and $\underline{b}$ are orthogonal or perpendicular and write $\underline{a} \perp \underline{b}$ if $(\underline{a}, \underline{b}) = 0$. The length of $\underline{a}$ is given by $\|\underline{a}\| = \sqrt{(\underline{a}, \underline{a})}$, the angle between non-zero vectors by $(\underline{a}, \underline{b}) = \|\underline{a}\| \|\underline{b}\| \cos \angle(\underline{a}, \underline{b})$.
More generally we say that V together with a map $(.,.) : V \times V \to \mathbb{R}$, is a real inner product space, if the following properties are satisfied:

(i) $(\underline{a}, \underline{b}) = (\underline{b}, \underline{a})$ symmetric;
(ii) $(\underline{a} + \underline{b}, \underline{c}) = (\underline{a}, \underline{c}) + (\underline{b}, \underline{c})$;
(iii) $(\lambda \cdot \underline{a}, \underline{b}) = \lambda \cdot (\underline{a}, \underline{b})$; (ii) & (iii): linear;
(iv) $(\underline{a}, \underline{a}) \geq 0$ with equality only for $\underline{a} = \underline{0}$ positive definite.

Of great importance is the theorem of Cauchy-Schwarz:

$$(\underline{a}, \underline{b})^2 \leq (\underline{a}, \underline{a})(\underline{b}, \underline{b}),$$

with equality if and only if the set $\{\underline{a}, \underline{b}\}$ is dependent (so $\underline{a} = \underline{0}$, or $\underline{b} = \lambda \underline{a}$ for some λ).

B.2.7 (orthonormal basis) A basis α for a real inner product space V is called *orthonormal* if $(\underline{a}_i, \underline{a}_j) = 0$ if $i \neq j$ and 1 if $i = j$, so the basis vectors are mutually orthogonal unit vectors. If $\{\underline{a}_1, \ldots, \underline{a}_m\}$ is an orthonormal basis of a subspace W, then the orthogonal projection of a vector $\underline{v} \in V$ on the subspace W is given by $\mathcal{P}_W(\underline{v}) = (\underline{v}, \underline{a}_1)\underline{a}_1 + \cdots + (\underline{v}, \underline{a}_m)\underline{a}_m$.

B.2.8 (Gram–Schmidt) Given a basis $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$ for a real inner product space V, an orthonormal basis β can be computed efficiently as follows:
Let $B_i =< \underline{b}_1, \ldots, \underline{b}_i >$ (so $B_0 = \{\underline{0}\}$). For $i = 1, \ldots, n$ define
$\underline{v}_i = \underline{a}_i - \mathcal{P}_{B_i}(\underline{a}_i)$ and then $\underline{b}_i = \underline{v}_i / \|\underline{v}_i\|$.
In practise it is more reasonable to work with the projection formula for mutually orthogonal basis vectors and only normalise in the end, so: $\underline{v}_1 = \underline{a}_1$, $\underline{v}_2 = \underline{a}_2 - \dfrac{(\underline{a}_2, \underline{v}_1)}{(\underline{v}_1, \underline{v}_1)}\underline{v}_1$, $\underline{v}_3 = \underline{a}_3 - (\ )\underline{v}_1 - (\ )\underline{v}_2$, etc..

B.3 Matrices

B.3.1 (matrices) An $m \times n$ matrix A is a rectangular block with m rows and n columns of numbers $a_{ij} \in \mathbb{K}$.

Abstract and more general: if R and C are sets, then an $R \times C$ matrix A is a map $R \times C \to \mathbb{K}$. Note the little difference, in the first definition, rows and colums are ordered.

B.3.2 (addition and scalar multiplication) If A and B are $m \times n$ matrices, then $S = A + B$ is the matrix with $s_{ij} = a_{ij} + b_{ij}$.
Abstract $S(r, c) = A(r, c) + B(r, c)$ for all $r \in R$ and $c \in C$.
If A is an $m \times n$ matrix, then λA is the matrix with ij-entry λa_{ij}.
Together with the zero matrix this turns the set of $m \times n$ matrices into a vector space, $\mathrm{Mat}_{m,n}(\mathbb{K})$ or $\mathrm{Mat}_{\mathbb{K}}(m, n)$ or also $\mathbb{K}^{m \times n}$.

B.3.3 (matrix multiplication) If A is an $m \times k$ matrix, and B is a $k \times n$ matrix, then the product $P = A \cdot B$ is the $m \times n$ matrix defined by

$$p_{ij} = \sum_{*=1}^{k} a_{i*} b_{*j},$$

More general: if A is an $R \times K$ and B a $K \times C$ matrix, then P is the $R \times C$ matrix defined by

$$p(r, c) = \sum_{* \in K} a(r, *) b(*, c).$$

Matrix multiplication is *associative*: $(AB)C = A(BC)$.
Proof: $ABC(i, j) = \sum_{kl} a_{ik} b_{kl} c_{li}$.

B.3.4 (vector vs matrix) We usually consider a vector $\underline{v} \in \mathbb{K}^n$ as a column vector, so we identify it with an $n \times 1$ matrix, again denoted $\underline{v}$. In case we want to view it as a row vector, or better a $1 \times n$ matrix, we write $\underline{v}^\top$. In this way we can define the products $\underline{u}^\top A$ and $A\underline{v}$ for $\underline{u} \in \mathbb{K}^m$, $\underline{v} \in \mathbb{K}^n$ and A an $m \times n$ matrix.
So, $\underline{v} = (x, y) \in \mathbb{K}^2$, corresponds to the matrices $\underline{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $\underline{v}^\top = (x \;\; y)$.
Or to make things ultimately clear: $(v_1, \ldots, v_n)^\top = (v_1 \; v_2 \; \cdots \; v_n)$.

Bibliography

[1] Carl B. Boyer. *A history of mathematics*. John Wiley & Sons, Inc., New York, (1989, 2nd edition with Uta C. Merzback)

[2] Bruce Cooperstein. *Elementary Linear Algebra: Methods, Procedures and Algorithms*. As suggested by the title: this book contains the description of algorithms with examples. See http://linearalgebramethods.com for more information. It is intended as a book for the theory, but very useful because of the examples.

[3] Jan van de Craats. *Vectoren en matrices*. Epsilon Uitgaven, Utrecht (2000).

This book is in dutch. Dit boek overlapt voor een deel met de bij het college besproken stof. De auteur zegt het soms net wat anders.

[4] David C. Lay. *Linear algebra and its applications*. Pearson/Addison Wesley, Boston, etc. (2006, 3rd ed. update) Nicely readable book, american style. Many exercises.

[5] Paul Halmos. *Finite–dimensional vector spaces*. Van Nostrand (1958) This book emphasizes the abstract theory of vector spaces.

[6] Murray R. Spiegel. *Theory and problems of complex variables*. Schaum's Outline Series, McGraw-Hill (1974)

[7] Hans Sterk. *Wat? Nog meer getallen*. Syllabus complexe getallen ten behoeve van Wiskunde D. Zie http://www.win.tue.nl/wiskunded (in Dutch).

[8] Gilbert Strang. *Linear algebra and its applications*. Harcourt Brace etc., San Diego (1988, 3rd edition)

Clearly en enthousiastically written text about linear algebra, with the emphasis on matrices. Contrary to Halmos' book is not chosen for the abstract approach.