Vector Spaces

7 Vector Spaces

7.1 Introduction

Preliminary remarks:

Definition 7.1 (*Group*) *Let G be a set and $*: G \times G \to G$ a function, so that*

G1 *Associativity:* $\forall g_1, g_2, g_3 \in G: \; g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$

G2 *Neutral element:* $\exists e \in G \; \forall g \in G: \; g * e = g$

G3 *Inverse element:* $\forall g \in G \; \exists_1 g^{-1} \in G: \; g * g^{-1} = e$

Then $(G, *)$ is called ***group****.*

If in addition

G4 *Commutativity:* $\forall g_1, g_2 \in G: \; g_1 * g_2 = g_2 * g_1$,

then it is called a ***commutative/abelian group****.*

We now add more structure by abstracting the familiar properties of real numbers with summation (subtraction) and multiplication (division).

Definition 7.3 (*Field*) *Let F be a set and $+ \colon K \times K \to K$ and $\cdot \colon K \times K \to K$ two functions such that*

F1 $(F, +)$ *is a commutative group (with neutral element 0)*

F2 $(F \setminus \{0\}, \cdot)$ *is a commutative group (with neutral element 1)*

F3 *Distributivity (compatibility of $+$ and $\cdot$)*

$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(a + b) \cdot c = a \cdot c + b \cdot c$$

We now abstract the notion of vectors in $\mathbb{R}^n$ and their properties.

Definition 7.5 (*Vector space*) *Let $\mathbb{F}$ be a field. A set V together with a function $+$ (sum) and a function $\cdot$ (scalar multiplication) with*

$$+ : V \times V \to V \qquad \cdot : \mathbb{F} \times V \to V$$
$$(v, w) \mapsto v + w \qquad (\lambda, v) \mapsto \lambda \cdot v$$

*is called **vector space (or linear space)** over $\mathbb{F}$, if the following axioms **VR1** and **VR2** hold:*

VR1 $(V, +)$ *is a commutative group with neutral element* 0.

VR2 *The scalar multiplication is compatible with $(V, +)$ in the following way: for $\lambda, \mu \in \mathbb{F}$, $v, w \in V$ it holds that*

i) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$
ii) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$
iii) $\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$
iv) $1 \cdot v = v$

Remarks:

- The vector space axioms allow for an abstract study and serve as an interface to the developed theory. For example, this is important for the study of linear equations, where the sought after solutions are functions (e.g., differential or integral equations). We then look for solutions in particular function spaces (typically infinite–dimensional which we approximate with finite–dimensional ones on the computer).
- Often one equips vector spaces with additional structure: Norm (abstract notion of length), inner product (relates to angles and orthogonality), topology (relates to limits, continuity and connectedness),... $\to$ Below we will introduce inner product spaces (a preliminary stage to so-called Hilbert Spaces)

7.2 Revisit: Linear Combination, Linear Independence, Basis

Based on the more abstract functions $+$ and $\cdot$ we can generally define:

Definition 7.7 *Let V be a vector space over the field $\mathbb{F}$ and $v_1, ..., v_r \in V$. Then, we define*

a) ***Linear combination:***

$$\sum_{j=1}^{r} \lambda_j v_j \quad \in V, \qquad \lambda_j \in \mathbb{F}$$

b) ***Span (set of all linear combinations):***

$$\mathit{span}(v_1, ..., v_r) := \{\sum_{j=1}^{r} \lambda_j v_j : \quad \lambda_j \in \mathbb{F}\,,\ j = 1, ..., r\} \in V$$

d) *The vectors $v_1, ..., v_r$ are called* **linearly independent**, *if*

$$\sum_{j=1}^{r} \lambda_j v_j = 0 \quad \Rightarrow \quad \lambda_j = 0\,,\ \forall j = 1, ..., r$$

e) *The vectors $v_1, ..., v_r$ are called basis of V, if*
   i) *$v_1, ..., v_r$ are linearly independent,*
   ii) *$\mathit{span}(v_1, ..., v_r) = V$.*

With the same proof as for $\mathbb{F}^n$ we can show:

Corollary 7.8 *For vectors $v_1, ..., v_r \in V$ the following statements are equivalent:*

i) *$v_1, ..., v_r$ are linearly independent*

ii) *every vector $v \in \mathit{span}(v_1, ..., v_r)$ can be uniquely linearly combined from the set $\{v_1, ..., v_r\}$.*

iii) *None of v_i for $i = 1, \ldots, r$ can be written as a linear combination of the other.*

Remark: Note that these notions coincide with the ones from the linear algebra section. However, now "vectors" are elements from any vector space V and could thus be vectors in the discrete sense, or matrices, or functions,...

Example 7.9

7.3 Linear functions: kernel, image, matrix representation

Definition 7.10 (*Linear function*) *Let V, W be two vector spaces over $\mathbb{F}$. Then a function $f : V \to W$ is called an $\mathbb{F}$-**linear function**, if*

$$f(\lambda v_1 +_V v_2) \;=\; \lambda f(v_1) +_W f(v_2) \quad \forall v_1, v_2 \in V,\ \lambda \in \mathbb{F}.$$

The set of all linear functions is denoted by $Hom_{\mathbb{F}}(V, W)$ (homomorphisms).
*For $f \in Hom_{\mathbb{F}}(V, W)$ we define the **kernel** $\ker(f)$ and the **image** $\mathrm{Im}(f)$ as*

$$\ker(f) := \{v \in V\colon\ f(v) = 0\} \subset V,$$

$$\mathrm{Im}(f) := \{f(v) \in W\colon\ v \in V\} \subset W.$$

Matrix Representation of Linear Mappings

Next we show that for finite–dimensional vector spaces V and W, say $\dim(V) = n$ and $\dim(W) = m$, the set of all linear functions from V to W is equivalent to the set of all $m \times n$ matrices.

Remark 7.12

- The matrix $\mathcal{M}_{\{w_1,\dots,w_m\}}^{\{v_1,\dots,v_n\}}(f)$ is sometimes called *transformation matrix* (in German: *Darstellungsmatrix*).
- It is the matrix representation of f under the given bases and it tells us how to transform the coordinates!
- Also, it establishes a one-to-one relation between $\mathrm{Hom}_{\mathbb{F}}(V, W)$ and $\mathbb{F}^{m\times n}$, i.e.,

$$\mathrm{Hom}_{\mathbb{F}}(V, W) \simeq \mathbb{F}^{m\times n}.$$

7.4 Inner Product Spaces

We now add more structure to a vector space: We equip it with an inner product.

Definition 7.14 *Let V be a vector space over $\mathbb{F}$.*

- *A mapping $(\cdot,\cdot)\colon V \times V \to \mathbb{F}$ is called **inner product** on $\mathbb{F}^n$ if it satisfies:*
  i) *Hermitian: $\forall x, y \in V:\ (x,y) = \overline{(y,x)}$,*
  ii) *Linear in its second argument: $\forall x, y_1, y_2 \in V, \lambda \in \mathbb{F}:\ (x, y_1 + \lambda y_2) = (x, y_1) + \lambda(x, y_2)$,*
  iii) *Positive definite: $\forall x \in V\backslash\{0\}:\ (x,x) > 0$.*
- *We call $(V, (\cdot,\cdot))$ an **Euclidean vector space** or **inner product space**.*

For simplicity we now stick to $V = \mathbb{R}^n$. We can show the following characterization:

Theorem 7.15 *Let $(\cdot,\cdot)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, then*

$$(\cdot,\cdot) \text{ inner product on } \mathbb{R}^n \;\Leftrightarrow\; \exists A \in \mathbb{R}^{n\times n}_{spd}\ \forall x, y \in \mathbb{R}^n : (x,y) := (x,y)_A := (x, Ay)_2.$$

In words: There is a one-to-one relation between all inner products on $\mathbb{R}^n$ and all symmetric and positive definite matrices in $\mathbb{R}^{n\times n}$. The proof is constructive and reveals that the entries of A are given by

$$a_{ij} := (e_i, e_j)$$

where e_j denote the standard basis vectors.

The important relation to Data Science/Optimization

- The gradient (times (-1)) of a function determines a descent direction. Moving in this direction will minimize the function
- The gradient is precisely the vector representing the derivative of a function.
- You will learn how to accelerate the gradient method by representing it in another inner product, e.g., the one induced by the Hessian $A = H_f(x)$ (this will give you Newton type method!).