Vector Spaces

7 Vector Spaces

7.1 Introduction

Preliminary remarks:

- So far, we worked with "vectors" in $\mathbb{R}^n, \mathbb{F}^n, \ldots$ We introduced the notions of summation, multiplication, linear combination, span, basis,...
- We also summed and scaled matrices, but did not talk about a span of matrices, basis, etc...
- In the exercise class we considered infinite discrete signals. We could add and scale sequences in $\mathbb{R}^{\mathbb{N}}, \mathbb{R}^{\mathbb{Z}}, ..$ as well and talk about linear combinations, basis, etc...
- Similarly, considering functions $f\colon \mathbb{R} \to \mathbb{R}$ and sets of functions (function spaces are studied in functional analysis). We could define summation and scaling. Example: Polynomials, frequencies,...

We now establish a more abstract point of view and revisit notions such as vectors, linear combinations, basis, etc. once again. You will see that we have already learned all the basic ideas.

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****.*

Group theory plays a crucial role in cryptography (RSA,...)

Example 7.2

- $G = \{g\}$
- $(\mathbb{Z}, +)$
- $(\mathrm{GL}(n, \mathbb{R}), \cdot)$; it is not commutative, take e.g.,

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

  In words: If both coordinates are scaled differently, then it makes a difference if we swap them before or after the scaling.

Not a group:

- $(\mathbb{N}, +)$, inverse element of 1?

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 $\cdot$)*

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

Example 7.4

- $(\mathbb{R}, +, \cdot)$
- $(\mathbb{Q}, +, \cdot)$
- $(\mathbb{C}, +, \cdot)$

Not a field:

- $(\mathbb{Z}, +, \cdot)$, because $(\mathbb{Z} \setminus \{0\}, \cdot)$ is not a group (no multiplicative inverse of, e.g., $2, 3, 4$...)
  It is just a "ring" (with respect to multiplication it is a semigroup (no inverses) and not a group)

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),...
  $\rightarrow$ Below we will introduce inner product spaces (a preliminary stage to so-called Hilbert Spaces)

Example 7.6

(i) $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$, C^n, $\mathbb{R}^{m\times n}$

Let us verify the axioms for $\mathbb{R}^{m\times n}$. We have defined matrix sum $+$ and scaling $\cdot$.

(VR1) $(\mathbb{R}^{m\times n}, +)$ is a commutative group:
...

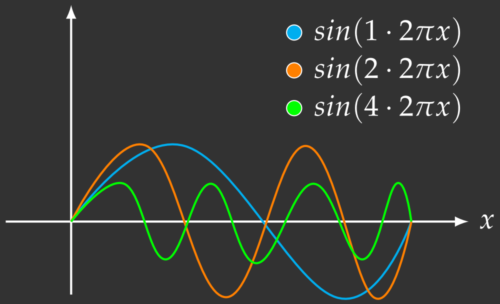
(VR2) Recall the compatibility properties from Lemma 1.5:
Let $A, B \in \mathbb{R}^{m\times n}$ and $r, s \in \mathbb{R}$. Then

$$\begin{array}{rc} i) & (r+s)\cdot A = r\cdot A + s\cdot A \\ ii) & r\cdot(A+B) = r\cdot A + r\cdot B \\ iii) & (r\cdot s)\cdot A = r\cdot(s\cdot A) \\ iv) & 1\cdot A = A \end{array}$$

(ii) Let $V := \{\rho \mid \rho : [0,1] \to \mathbb{R}\}$ and define

$$(\lambda \cdot \rho)(x) := \lambda \cdot \rho_2(x), \qquad (\rho_1 + \rho_2)(x) := \rho_1(x) + \rho_2(x)$$

Functional analysis deals with infinite dimensional vector spaces of functions and generalizes/extends results from linear algebra.



(iii) Gray-scale images of size 1024 x 1024:

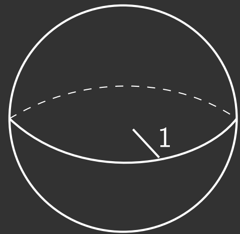
$$\{A \in \mathbb{R}^{1024\times 1024} \mid a_{ij} \in [0,1]\}$$

is only a subset of a vector space, not a vector space as such. It is a convex set though.

(iv) The two-dimensional sphere in $\mathbb{R}^3$ is defined by

$$S^2 = \{x \in \mathbb{R}^3 \mid \|x\| = 1\}$$

for some norm $\|\cdot\|$. For the Euclidean norm it looks like this:



This not a vector space (curvature prevents a simple summation of two elements), but a so-called **manifold**.

(v) Let $n \in \mathbb{N}$ and $P_n(\mathbb{R})$ be the set of all polynomials of degree $\leq n$ on $\mathbb{R}$, i.e., the set $P_n(\mathbb{R})$ contains all functions $p : \mathbb{R} \to \mathbb{R}$ of the form

$$p(x) = \sum_{k=0}^{n} \alpha_k x^k$$

for some $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$. We define a summation and scalar multiplication:

$$\begin{aligned} +&: P_n(\mathbb{R}) \times P_n(\mathbb{R}) \to P_n(\mathbb{R}), \ (p+q)(x) := p(x) + q(x), \\ \cdot&: \mathbb{R} \times P_n(\mathbb{R}) \to P_n(\mathbb{R}), \qquad (r \cdot p)(x) := r \cdot p(x). \end{aligned}$$

Monomials $q_k : \mathbb{R} \to \mathbb{R}, \ x \mapsto x^k$ form a basis of this $(n+1)$–dimensional vector space.

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

Let $V := \mathbb{R}^{2\times 3}$, then a basis of length 6 is given by

$$\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

Independent: ✓

Generating:

$$\begin{aligned} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} &= a_{11} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{12} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{13} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ &+ a_{21} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{22} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + a_{23} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

The basis is of length 6 and thus $\dim(V) = 6 (= 2 \cdot 3)$.

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 $\mathrm{Hom}_{\mathbb{F}}(V, W)$ (homomorphisms).
*For $f \in \mathrm{Hom}_{\mathbb{F}}(V, W)$ we define the **kernel** $\ker(f)$ and the **image** $\mathrm{Im}(f)$ as*

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

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

Example 7.11

(i) $0 : V \to W,\ v \mapsto 0 \in W$

Check: $\quad 0(\lambda v_1 + v_2) = 0 = 0 + 0 = \lambda 0(v_1) + 0(v_2) \quad \forall v_1, v_2 \in V,\ \lambda \in \mathbb{F}.$

(ii) $\mathrm{id}\colon V \to V,\ v \mapsto v$

Check: $\quad \mathrm{id}(\lambda v_1 + v_2) = \lambda v_1 + v_2 = \lambda \mathrm{id}(v_1) + \mathrm{id}(v_2) \quad \forall v_1, v_2 \in V,\ \lambda \in \mathbb{F}.$

(iii) If $\{v_1, \dots, v_n\}$ is a basis of V and $\lambda \in \mathbb{F}^n$ the coordinate representation of $v \in V$, i.e., $v = \sum_{i=1}^n \lambda_i v_i$. Then, the function

$$\pi_i : V \to \mathbb{F}, \quad v \mapsto \lambda_i$$

is linear.

(iv) The derivative $\frac{d}{dx}$ and the integral $\int$ operators are linear functions.

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.

Introductory example:
Let us consider polynomials of degree no higher than 2, i.e., $\mathcal{P}_2(\mathbb{R})$. Then an example of a linear function from $\mathcal{P}_2(\mathbb{R})$ $(n = 3)$ to $\mathcal{P}_1(\mathbb{R})$ $(m = 2)$ is given by the derivative $f(\cdot) := \frac{d}{dx}(\cdot)$.

Examples:

- $p(x) = x + 2x^2 \quad \simeq (0, 1, 2)$
  $\frac{d}{dx}p(x) = 1 + 4x \quad \simeq (1, 4)$
- $p(x) = 1 + 3x^2 \quad \simeq (1, 0, 3)$
  $\frac{d}{dx}p(x) = 6x \quad \simeq (0, 6)$

In terms of the coordinates, the derivative $\frac{d}{dx}$ can be represented by the matrix

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

[ex:commutative diagram]

In general

Let V and W be finite–dimensional vector spaces over $\mathbb{F}$ with $\dim(V) = n$ and $\dim(W) = m$ and bases $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$, respectively.

Then for all $1 \leq j \leq n$ we can represent the $f(v_j) \in W$ in the basis of W, i.e., for some numbers $(a_{ij})_{1 \leq i \leq m} \in \mathbb{F}^m$ we find

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Given these fixed bases for V and W, we define the matrix

$$\mathcal{M}^{\{v_1,\ldots,v_n\}}_{\{w_1,\ldots,w_m\}}(f) := (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathbb{F}^{m \times n}.$$

Since any $v \in V$ can be written as

$$v = \sum_{j=1}^{n} \lambda_j v_j,$$

for some $\lambda_j \in \mathbb{F}$, we find by linearity of f and associativity that

$$f(v) = f(\sum_{j=1}^{n} \lambda_j v_j) = \sum_{j=1}^{n} \lambda_j f(v_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \lambda_j a_{ij} \right) w_i = \sum_{i=1}^{m} \left(\mathcal{M}^{\{v_1,\ldots,v_n\}}_{\{w_1,\ldots,w_m\}}(f) \cdot \lambda \right)_i w_i. \tag{17}$$

Modulo the representation in the bases we see that at the core of the evaluation $f(v)$ the matrix vector product

$$\mathcal{M}^{\{v_1,\ldots,v_n\}}_{\{w_1,\ldots,w_m\}}(f) \cdot \lambda,$$

which translates the coordinates accordingly.

[ex:Draw commutative diagram.]

Remark 7.12

- The matrix $\mathcal{M}^{\{v_1,\dots,v_n\}}_{\{w_1,\dots,w_m\}}(f)$ 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.,

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

Example 7.13

i) Exercise: Let $f \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F},\mathbb{F})$ and $v_1 \neq 0 \neq w_1$ two real numbers, each forming a basis of $\mathbb{F}$.

ii) Let $f \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n,\mathbb{F}^m)$ and consider the two bases $V = (v_1,\dots,v_n) \in \mathrm{GL}_n(\mathbb{F})$, and $W = (w_1,\dots,w_m) \in \mathrm{GL}_m(\mathbb{F})$. Note that the basis vectors are now vectors in our sense so far and we can now stack them into matrices. Also, for simplicity we write

$$\mathcal{M}^V_W(f) := \mathcal{M}^{\{v_1,\dots,v_n\}}_{\{w_1,\dots,w_m\}}(f).$$

We observe

$$f(v_j) = \sum_{j=1}^{m} a_{ij} w_j \quad \Leftrightarrow \quad (a_{ij})_i = W^{-1} f(v_j),$$

so that the representing matrix of f is given by

$$\mathcal{M}^V_W(f) := W^{-1}(f(v_1),\dots,f(v_n)) \in \mathbb{F}^{m\times n}$$

and

$$f(v) = f(VV^{-1}v) = W\mathcal{M}^V_W(f)V^{-1}v.$$

[ex: Draw commutative diagram.]

iia) Take for example $f(v) := Av$ for some $A \in \mathbb{F}^{m \times n}$, so that $f \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m)$. Then

$$\mathcal{M}_I^I(f) = ... = A$$

and

$$\mathcal{M}_W^V(f) = ... = W^{-1}AV$$

iib) In particular, if $m = n$ (so we can choose the same basis for $V = W = \mathbb{F}^n$), we observe

$$\mathcal{M}_I^I(f) = ... = A$$

and

$$\mathcal{M}_V^V(f) = ... = V^{-1}AV.$$

Observations:

- $\mathcal{M}_I^I(f)$ and $\mathcal{M}_V^V(f)$ are similar. Hence, they have the same eigenvalues.
- The matrix representation depends on the chosen basis, but the somewhat deeper properties (eigenvalues,...) of the linear map as such are unchanged.
- This leads to the idea of searching for a particular basis in order to get a matrix representation, which reveals the "essence" of the respective linear function. If the eigenvectors form a basis then they would be a good choice: In this case the matrix representation is diagonal having precisely the eigenvalues on its diagonal.

iii) Consider

$$\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}) = \{f\colon \mathbb{R}^n \to \mathbb{R} : f \text{ linear}\} \quad \text{(dual space of } \mathbb{R}^n\text{)}.$$

Let $V = (v_1, \ldots, v_n) \in \mathrm{GL}_n(\mathbb{R})$ and $\{1\}$ be a bases for V and $\mathbb{R}$, respectively. Then for $f \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ obviously $f(v_j) = f(v_j) \cdot 1$, so that

$$\mathcal{M}_I^V(f) = \mathcal{M}_{\{1\}}^{\{v_1,\ldots,v_n\}}(f) = (f(v_1), \ldots, f(v_n)) \in \mathbb{R}^{1\times n}$$

is simply a row vector.

For $v = VV^{-1}v =: V\lambda \in \mathbb{R}^n$ we have

$$f(v) = \sum_{j=1}^{n} \lambda_j f(v_j) = \mathcal{M}_{\{1\}}^{\{v_1,\ldots,v_n\}}(f) \cdot \lambda = \mathcal{M}_{\{1\}}^{\{v_1,\ldots,v_n\}}(f) V^{-1} \cdot v = \left(\left(\mathcal{M}_{\{1\}}^{\{v_1,\ldots,v_n\}}(f) V^{-1} \right)^\top, v \right)_2.$$

Important observations/remarks

- In words: The column vector $v_f := \left(\mathcal{M}_{\{1\}}^{\{v_1,\ldots,v_n\}}(f) V^{-1} \right)^\top$ represents the linear function f in the standard inner product

$$f(\cdot) = (v_f, \cdot)_2.$$

  This is the finite–dimensional "light version" of the Riesz representation theorem.
- If we choose the standard basis $V = I$, then this vector is simply

$$\mathcal{M}_I^I(f)^\top = \begin{pmatrix} f(v_1) \\ \vdots \\ f(v_n) \end{pmatrix} \in \mathbb{R}^n.$$

- In the next section we consider general inner products $(\cdot,\cdot)_A$ on $\mathbb{R}^n$ and how this effects the representing vector v_f:

$$f(\cdot) = (v_f^A, \cdot)_A = (v_f, \cdot)_2.$$

- When considering the derivative $DF(x)\colon \mathbb{R}^n \to \mathbb{R}$ as a linear mapping we will call the vector $\mathcal{M}_I^I(DF(x))^\top$ the gradient of F at point x (represented in the standard basis and standard inner product).

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.

Final note:

Let us again consider $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ and now a general inner product $(\cdot,\cdot)_A$ on $\mathbb{R}^n$ for some spd matrix $A \in \mathbb{R}^{n\times n}$.

Question

For $f \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ we know $f\colon \mathbb{R}^n \to \mathbb{R}$ is linear. On the other hand $(v,\cdot)_A : \mathbb{R}^n \to \mathbb{R}$ is linear for a fixed vector $v \in \mathbb{R}^n$ and inner product $(\cdot,\cdot)_A$, i.e., $(v,\cdot)_A \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$. Thus it is natural to ask whether there is a one-to-one relation between $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ and $\mathbb{R}^n$, i.e., can we find a unique $v_f^A \in \mathbb{R}^n$ for $f \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ so that $f(\cdot) = (v_f^A, \cdot)_A$?

Yes we can: In fact, define $v_f := v_f^I := \left(\mathcal{M}_I^I(f)\right)^\top$, then using the findings from above and the properties of A we get

$$f(v) = \left(v_f, v\right)_2 = v_f^\top v = v_f^\top (A^{-1}A)v = v_f^\top (A^{-1})^\top A v = (A^{-1}v_f, Av)_2 = (A^{-1}v_f, v)_A.$$

In words:

- Fixing to the standard inner product $A = I$, we find

$$v_f^I = \left(\mathcal{M}_I^I(f)\right)^\top$$

- Changing to a general inner product, we find

$$v_f^A = A^{-1}v_f^I$$

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!).