**EE951: Linear Algebra** **eMasters**

# Chapter 9: Linear Functions

*Author: Ketan Rajawat*

**Disclaimer**: *These notes are taken from various sources, sometimes without proper citation. Do not distribute outside the class or upload on the Internet.*

## Contents

**9.1 Linear Functions** **9-1**

**9.2 Linear Functions and Superposition** **9-1**

**9.3 Vector-valued Linear Functions** **9-4**

## 9.1 Linear Functions

We will commonly work with functions that take in $n$ inputs (collected in an $n$-dimensional real vector) and output real scalars. These are denoted as

$$f : \mathbb{R}^n \to \mathbb{R}$$

Here, $f$ is the name of the function and the arrow $\to$ signifies the mapping from the domain to the codomain. The domain of a function refers to the set of all possible inputs, in this case, $n$-dimensional real vectors. In our notation, the domain is $\mathbb{R}^n$. The codomain of a function represents the set of all possible outputs. In this case, the codomain is the set of real numbers, denoted as $\mathbb{R}$. Such a function is also called a scalar function or a scalar-valued function, to differentiate it from more general vector-valued functions denoted by $h : \mathbb{R}^n \to \mathbb{R}^m$.

For example, let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as $f(\mathbf{x}) = x_1 + 2x_2$, where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is a 2-dimensional vector. Here, the domain of the function is all 2-dimensional real vectors, and the codomain is the set of real numbers. The function takes in a vector $\mathbf{x}$ and outputs a scalar by adding the first component $x_1$ to twice the second component $2x_2$.

**Example 9.1.** Consider the **inner product function**. Given $\mathbf{a} \in \mathbb{R}^n$, let us consider the function $f : \mathbb{R}^n \to \mathbb{R}$ defined as $f(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x} \rangle = a_1x_1 + a_2x_2 + \ldots + a_nx_n$, for all $\mathbf{x} \in \mathbb{R}^n$. In this example, the domain of the function is $\mathbb{R}^n$, representing all $n$-dimensional real vectors. The function takes in a vector $\mathbf{x}$ and forms a weighted sum of its entries using the weights specified by $\mathbf{a}$. Each entry of $\mathbf{a}$, $a_i$, corresponds to the weight applied to the corresponding entry $x_i$ of $\mathbf{x}$. The inner product function calculates the dot product between $\mathbf{a}$ and $\mathbf{x}$, resulting in a scalar output.

## 9.2 Linear Functions and Superposition

A linear function, also known as a linear transformation or linear map, is a function between vector spaces that preserves vector addition and scalar multiplication. In other words, a linear function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies two key properties:

1. *Additivity:* For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ in the domain of the function, the function applied to the sum of $\mathbf{u}$ and $\mathbf{v}$ is equal to the sum of the function applied to $\mathbf{u}$ and the function applied to $\mathbf{v}$. Mathematically, this can be expressed as: $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$

2. *Homogeneity:* For any vector $\mathbf{u} \in \mathbb{R}^n$ in the domain of the function and any scalar $\alpha \in \mathbb{R}$, the function applied to the scalar multiple of $\mathbf{u}$ is equal to the scalar multiple of the function applied to $\mathbf{u}$. Mathematically, this can be expressed as: $f(\alpha\mathbf{u}) = \alpha f(\mathbf{u})$

The two properties can be combined into a single property called superposition, which refers to the ability to express a function's action on a linear combination of vectors by examining its action on each individual vector and scaling the results accordingly. Mathematically, if we have a linear function $f$, vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in the domain of $f$, and scalars $\alpha_1$ and $\alpha_2$, the superposition property states:

$$f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) \tag{9.1}$$

Observe that in this inequality, the left-hand side involves scaling of vectors while the right-hand side involves scaling of the function values.

**Example 9.2.** Let us show that the inner product function $f(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x} \rangle = a_1x_1 + a_2x_2 + \ldots + a_nx_n$ satisfies the superposition equality. To do this, we need to demonstrate that it holds for any vectors $\mathbf{x}$ and $\mathbf{y}$ in the domain and scalars $\alpha$ and $\beta$.

Let us consider:

$$\begin{aligned}
f(\alpha\mathbf{x} + \beta\mathbf{y}) &= \langle \mathbf{a}, \alpha\mathbf{x} + \beta\mathbf{y} \rangle && \text{definition of } f && (9.2)\\
&= \langle \mathbf{a}, \alpha\mathbf{x} \rangle + \langle \mathbf{a}, \beta\mathbf{y} \rangle && \text{distributivity with vector addition} && (9.3)\\
&= \alpha\langle \mathbf{a}, \mathbf{x} \rangle + \beta\langle \mathbf{a}, \mathbf{y} \rangle && \text{associativity with scalar multiplication} && (9.4)\\
&= \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) && \text{definition of } f && (9.5)
\end{aligned}$$

Thus, we have shown that the inner product function satisfies the superposition equality.

This result indicates that the inner product function is a linear function enabling us to express the action of the inner product function on a linear combination of vectors in terms of the corresponding linear combination of their inner products with $\mathbf{a}$.

This also implies that the following functions, which are expressible as inner products of constant vectors with $\mathbf{x}$, are also linear

1. $f(\mathbf{x}) = \frac{1}{n}\sum_i x_i$
2. $f(\mathbf{x}) = \sum_i x_i$
3. $f(\mathbf{x}) = x_i$ for some $1 \leq i \leq n$

**Example 9.3.** The function $g(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x} \rangle + b$ is not a linear function. To see this, let us consider two vectors $\mathbf{x}$ and $\mathbf{y}$ in the domain and scalars $c$ and $d$. We will examine whether the superposition property holds for $g(\mathbf{x})$ and $g(\mathbf{y})$:

$$\begin{aligned} g(c\mathbf{x}+d\mathbf{y}) &= \langle \mathbf{a}, c\mathbf{x}+d\mathbf{y}\rangle + b && \text{definition of } g && (9.6)\\ &= \langle \mathbf{a}, c\mathbf{x}\rangle + \langle \mathbf{a}, d\mathbf{y}\rangle + b && \text{distributivity with vector addition} && (9.7)\\ &= c\langle \mathbf{a}, \mathbf{x}\rangle + d\langle \mathbf{a}, \mathbf{y}\rangle + b && \text{associativity with scalar multiplication} && (9.8)\\ &\neq cg(\mathbf{x}) + dg(\mathbf{y}) && \text{unless } b = 0 && (9.9) \end{aligned}$$

In the last step, we can see that the superposition property does not hold. The terms $b$ in $g(c\mathbf{x}+d\mathbf{y})$ and $c(g(\mathbf{x})) + d(g(\mathbf{y}))$ are not equal unless $b = 0$. This means that for any value of $b$ other than 0, the superposition property fails.

Therefore, we can conclude that the function $g(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x}\rangle + b$ is not a linear function because it does not satisfy the superposition property for all vectors $\mathbf{x}$, $\mathbf{y}$ and scalars $c$, $d$.

We have seen that inner product is a linear function. Remarkably, the converse is also true, that is, any linear function can be expressed as an inner product. We have the following result.

*Lemma* 9.1. Any linear function can be expressed as an inner product.

*Proof:* We need to demonstrate that for any linear function $f : \mathbb{R}^n \to \mathbb{R}$, there exists a vector $\mathbf{a} \in \mathbb{R}^n$ such that $f(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x}\rangle$ for all $\mathbf{x} \in \mathbb{R}^n$.

Let us consider a linear function $f : \mathbb{R}^n \to \mathbb{R}$. By linearity, we know that for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and any scalar $c$, the following properties hold:

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \qquad f(c\mathbf{x}) = cf(\mathbf{x}) \tag{9.10}$$

Now, we define a vector $\mathbf{a} = \begin{bmatrix} f(\mathbf{e}_1)\\ f(\mathbf{e}_2)\\ \vdots \\ f(\mathbf{e}_n)\end{bmatrix}$, where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the standard unit vectors of $\mathbb{R}^n$.

Let us consider an arbitrary vector $\mathbf{x} \in \mathbb{R}^n$. Recall from earlier that we can express $\mathbf{x}$ as a linear combination of the standard unit vectors: $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$. Now, using the linearity of the function $f$, we can apply it to each term:

$$\begin{aligned} f(\mathbf{x}) &= f(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n)\\ &= x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + \ldots + x_n f(\mathbf{e}_n)\\ &= \langle \mathbf{a}, \mathbf{x}\rangle \end{aligned}$$

Thus, we have expressed the linear function $f(\mathbf{x})$ as an inner product $\langle \mathbf{a}, \mathbf{x}\rangle$ using the vector $\mathbf{a}$.

This result demonstrates that any linear function $f : \mathbb{R}^n \to \mathbb{R}$ can indeed be expressed as an inner product with a corresponding vector $\mathbf{a}$ in $\mathbb{R}^n$. The vector $\mathbf{a}$ captures the behavior of the linear function on the standard unit vectors, allowing us to represent the function using the elegant inner product notation. ∎

In particular, the formula

$$f(\mathbf{x}) = x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + \ldots + x_n f(\mathbf{e}_n) \tag{9.11}$$

holds for all linear functions, and provides an alternative way of evaluating $f$.

Suppose that $f$ is specified as a program or routine that given any input $\mathbf{x}$ returns $f(\mathbf{x})$. Then, if we know that $f$ is linear, we can use the formula in (9.11) to calculate $f(\mathbf{e}_1)$, $f(\mathbf{e}_2)$, …, $f(\mathbf{e}_n)$ and hence determine

$f$ entirely. In this setting, by precomputing and storing the $\{f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n)\}$, we can efficiently compute $f(\mathbf{x})$ using the formula (9.11) without needing to call the original routine for each value of $\mathbf{x}$.

The next result talks about the uniqueness of the inner product representation.

*Lemma* 9.2. The representation of a linear function as an inner product is unique.

*Proof:* We prove the desired result by the way of contradiction. Let us assume, contrary to the desired statement, that there are indeed two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ such that $f(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle$ and $f(\mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle$ for all $\mathbf{x} \in \mathbb{R}^n$.

Now, consider the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in $\mathbb{R}^n$. By substituting $\mathbf{x} = \mathbf{e}_i$ into the given representations of the linear function $f(\mathbf{x})$, we have:

$$\begin{aligned} f(\mathbf{e}_i) &= \langle \mathbf{u}, \mathbf{e}_i \rangle = u_i && \text{(from the first representation)} \quad (9.12) \\ &= \langle \mathbf{v}, \mathbf{e}_i \rangle = v_i && \text{(from the second representation)} \quad (9.13) \end{aligned}$$

This implies that for each $i = 1, 2, \ldots, n$, the $i$th component of $\mathbf{u}$ is equal to the $i$th component of $\mathbf{v}$, i.e., $u_i = v_i$ for all $i$ or equivalently $\mathbf{u} = \mathbf{v}$.

Therefore, we have shown that if a linear function $f : \mathbb{R}^n \to \mathbb{R}$ can be expressed as $f(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x} \rangle$, then this representation is unique, and there exists only one vector $\mathbf{a}$ that satisfies this condition. ■

Together, the two results can be combined into the Riesz Representation Theorem: for a linear function $f : \mathbb{R}^n \to \mathbb{R}$, there exists a unique vector $\mathbf{u} \in \mathbb{R}^n$ such that $f(\mathbf{v}) = \mathbf{u}^\mathsf{T}\mathbf{v}$ for every $\mathbf{v} \in \mathbb{R}^n$.

## 9.3 Vector-valued Linear Functions

So far we looked at scalar-valued functions, but we can similarly also define vector-valued functions of the form $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. In programming parlance, the function takes an $n$-vector as input and produces an $m$-vector as output.

A vector-valued linear function $\mathbf{f}$ also satisfies superposition in a similar way, i.e., given vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in the domain of $\mathbf{f}$, and scalars $\alpha_1$ and $\alpha_2$, the superposition property holds for linear functions:

$$\mathbf{f}(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) = \alpha_1 \mathbf{f}(\mathbf{v}_1) + \alpha_2 \mathbf{f}(\mathbf{v}_2) \tag{9.14}$$

As a simple example, consider the function $\mathbf{p} : \mathbb{R}^2 \to \mathbb{R}^2$ that projects a point on to the $x$-axis, i.e.,

$$\mathbf{p}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 0 \end{bmatrix}. \tag{9.15}$$

In other words, the function keeps the first element of the vector but makes the second element zero. In the two-dimensional plane, we call the coordinate $(x, 0)$ to be the projection of $(x, y)$ on to the x-axis. This is depicted in Fig. 9.1.

We can consider a vector valued function $\mathbf{f}(\mathbf{x})$ as being a stack of several scalar-valued functions, i.e.,

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} \tag{9.16}$$

Since the vector-valued function is linear, each of its element must also be a linear function, which from Lemma 9.1, can be expressed as an inner product, i.e,

$$f_i(\mathbf{x}) = \mathbf{a}_i^\mathsf{T}\mathbf{x} \tag{9.17}$$

Stacking these vectors together, we can write

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \mathbf{a}_1^\mathsf{T}\mathbf{x} \\ \mathbf{a}_2^\mathsf{T}\mathbf{x} \\ \vdots \\ \mathbf{a}_m^\mathsf{T}\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^\mathsf{T} \\ \mathbf{a}_2^\mathsf{T} \\ \vdots \\ \mathbf{a}_m^\mathsf{T} \end{bmatrix} \mathbf{x} = \mathbf{A}\mathbf{x} \tag{9.18}$$

where the matrix $\mathbf{A}$ has rows $\mathbf{a}_1^\mathsf{T}$, $\mathbf{a}_2^\mathsf{T}$, …, $\mathbf{a}_m^\mathsf{T}$. In other words, any vector-valued linear function can be expressed as a matrix-vector product.

**Example 9.4.** A simple example is the scaling function, where each coordinate is scaled by a positive scalar. The scaling function is given by

$$\mathbf{S}(\mathbf{x}) = \mathrm{Diag}\,(\mathbf{s})\,\mathbf{x} \tag{9.19}$$

where $\mathbf{s} \in \mathbb{R}^n$ is an $n$-vector containing positive real numbers. Specifically, we can see that

$$[\mathbf{S}(\mathbf{x})]_i = s_i x_i \qquad\qquad i \in \{1, 2, \ldots, n\}. \tag{9.20}$$

**Example 9.5.** Consider the following vector-valued function, that maps a point $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in three-dimensional space to a point $\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$ in two-dimensional space by forcing the last coordinate to be zero, i.e.,

$$\mathbf{p}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}. \tag{9.21}$$

The idea is also depicted in Fig. 9.1. Such a function is called the orthogonal projection, because the dashed line is orthogonal to the $x - y$ plane and the resulting point is a projection of the original three-dimensional point on a plane, much in the same way, as an image on a wall is a projection of an actual three-dimensional scene.

The corresponding projection matrix is given by

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so that $\mathbf{p}(\mathbf{x}) = \mathbf{P}\mathbf{x}$. An interesting property of such a projection matrix is that it is idempotent, i.e., $\mathbf{P}^2 = \mathbf{P}$. We see that

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

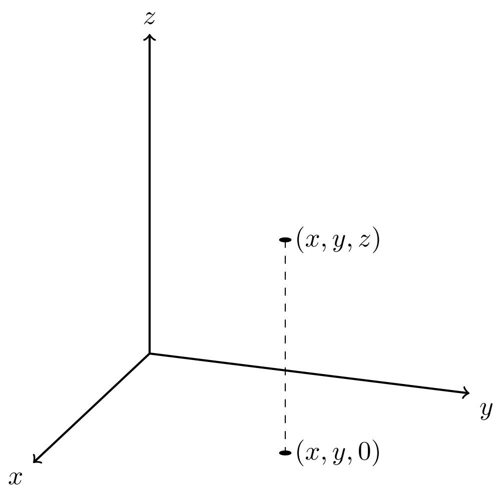


Figure 9.1: Projection of the point $(x, y, z)$ on to the $x - y$ plane.

which confirms that the projection matrix is idempotent.