Linear Algebra 1

Lecture #4

Linear functions

Irena Penev

October 30, 2023

Definition

For a field $\mathbb{F}$, a function $f : \mathbb{F}^m \to \mathbb{F}^n$ is said to be a *linear function* (or a *linear transformation*) if it satisfies the following two conditions (axioms):

1. for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^m$, we have that $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$;
2. for all vectors $\mathbf{u} \in \mathbb{F}^m$ and scalars $\alpha \in \mathbb{F}$, we have that $f(\alpha\mathbf{u}) = \alpha f(\mathbf{u})$.

Definition

For a field $\mathbb{F}$, a function $f : \mathbb{F}^m \to \mathbb{F}^n$ is said to be a *linear function* (or a *linear transformation*) if it satisfies the following two conditions (axioms):

1. for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^m$, we have that $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$;
2. for all vectors $\mathbf{u} \in \mathbb{F}^m$ and scalars $\alpha \in \mathbb{F}$, we have that $f(\alpha\mathbf{u}) = \alpha f(\mathbf{u})$.

Proposition 1.10.1

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then for all vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{F}^m$ and all scalars $\alpha_1, \dots, \alpha_k \in \mathbb{F}$, we have that

$$f\Big(\alpha_1\mathbf{v}_1 + \dots + \alpha_k\mathbf{v}_k\Big) \;=\; \alpha_1 f(\mathbf{v}_1) + \dots + \alpha_k f(\mathbf{v}_k).$$

Proof. Easy induction (details: exercise).

Example 1.10.2

Determine whether the following functions are linear (and prove your answer):

(a) $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{R}$.

(b) $g : \mathbb{Z}_2^2 \to \mathbb{Z}_2^4$ given by $g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 \\ 1 \end{bmatrix}$ for all $x_1, x_2 \in \mathbb{Z}_2$.

(c) $h : \mathbb{Z}_3^3 \to \mathbb{Z}_3^2$ given by $h\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{Z}_3$.

- **Remark:**
  - To show that a function is linear, we must show that it satisfies both axioms from the definition of a linear function; in particular, axiom 1 must hold for **all** vectors **u** and **v**, and axiom 2 must hold for **all** vectors **u** and scalars α.

- **Remark:**
  - To show that a function is linear, we must show that it satisfies both axioms from the definition of a linear function; in particular, axiom 1 must hold for **all** vectors $\mathbf{u}$ and $\mathbf{v}$, and axiom 2 must hold for **all** vectors $\mathbf{u}$ and scalars α.
  - On the other hand, to show that a function is **not** linear, it is enough to show that it fails to satisfy at least one of the axioms 1 and 2 from the definition of a linear function.

- **Remark:**
  - To show that a function is linear, we must show that it satisfies both axioms from the definition of a linear function; in particular, axiom 1 must hold for **all** vectors $\mathbf{u}$ and $\mathbf{v}$, and axiom 2 must hold for **all** vectors $\mathbf{u}$ and scalars α.
  - On the other hand, to show that a function is **not** linear, it is enough to show that it fails to satisfy at least one of the axioms 1 and 2 from the definition of a linear function.
  - To show that a function does **not** satisfy axiom 1, it is enough to exhibit **one particular pair of vectors** $\mathbf{u}$ and $\mathbf{v}$ for which that axiom does not hold.
  - Similarly, to show that a function does **not** satisfy axiom 2, it is enough to exhibit **one particular vector** $\mathbf{u}$ and **one particular scalar** α for which axiom 2 fails.

Example 1.10.2

(a) $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{R}$.

Solution.

Example 1.10.2

(a) $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{R}$.

Solution. (a) The function f is linear. We prove this by verifying the axioms of a linear function for the function f, as follows.

Example 1.10.2

(a) $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{R}$.

Solution. (a) The function f is linear. We prove this by verifying the axioms of a linear function for the function f, as follows.

1. Fix vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$. WTS $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$.

Example 1.10.2

(a) $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{R}$.

Solution. (a) The function f is linear. We prove this by verifying the axioms of a linear function for the function f, as follows.

1. Fix vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$. WTS $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$. For this, we compute (next slide):

Solution (continued).

$$
\begin{aligned}
f(\mathbf{u}+\mathbf{v}) &= f\left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}\right) = f\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix}\right) \\
&\overset{(*)}{=} \begin{bmatrix} (u_1+v_1)-(u_2+v_2)+(u_3+v_3) \\ (u_1+v_1)+(u_2+v_2) \end{bmatrix} \\
&= \begin{bmatrix} (u_1-u_2+u_3)+(v_1-v_2+v_3) \\ (u_1+u_2)+(v_1+v_2) \end{bmatrix} \\
&= \begin{bmatrix} u_1-u_2+u_3 \\ u_1+u_2 \end{bmatrix} + \begin{bmatrix} v_1-v_2+v_3 \\ v_1+v_2 \end{bmatrix} \\
&\overset{(**)}{=} f\left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\right) + f\left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}\right) \\
&= f(\mathbf{u}) + f(\mathbf{v}),
\end{aligned}
$$

where both (*) and (**) follow from the definition of f.

Solution (continued). Fix a vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ in $\mathbb{R}^3$ and a scalar $\alpha \in \mathbb{R}$. WTS $f(\alpha\mathbf{u}) = \alpha f(\mathbf{u})$. For this, we compute:

$$\begin{array}{rcl} f(\alpha\mathbf{u}) & = & f\Big(\alpha \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\Big) \;=\; f\Big(\begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \alpha u_3 \end{bmatrix}\Big) \\ & \overset{(*)}{=} & \begin{bmatrix} \alpha u_1 - \alpha u_2 + \alpha u_3 \\ \alpha u_1 + \alpha u_2 \end{bmatrix} \;=\; \begin{bmatrix} \alpha(u_1 - u_2 + u_3) \\ \alpha(u_1 + u_2) \end{bmatrix} \\ & = & \alpha \begin{bmatrix} u_1 - u_2 + u_3 \\ u_1 + u_2 \end{bmatrix} \;\overset{(**)}{=}\; \alpha f\Big(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\Big), \end{array}$$

where both (*) and (**) follow from the definition of f.

Solution (continued). Fix a vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ in $\mathbb{R}^3$ and a scalar $\alpha \in \mathbb{R}$. WTS $f(\alpha\mathbf{u}) = \alpha f(\mathbf{u})$. For this, we compute:

$$\begin{aligned}
f(\alpha\mathbf{u}) &= f\Big(\alpha \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\Big) = f\Big(\begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \alpha u_3 \end{bmatrix}\Big) \\
&\overset{(*)}{=} \begin{bmatrix} \alpha u_1 - \alpha u_2 + \alpha u_3 \\ \alpha u_1 + \alpha u_2 \end{bmatrix} = \begin{bmatrix} \alpha(u_1 - u_2 + u_3) \\ \alpha(u_1 + u_2) \end{bmatrix} \\
&= \alpha \begin{bmatrix} u_1 - u_2 + u_3 \\ u_1 + u_2 \end{bmatrix} \overset{(**)}{=} \alpha f\Big(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\Big),
\end{aligned}$$

where both (*) and (**) follow from the definition of f.

We have now shown that f satisfies both axioms from the definition of a linear function. So, f is linear, as we had claimed. $\square$

Example 1.10.2

(b) $g : \mathbb{Z}_2^2 \to \mathbb{Z}_2^4$ given by $g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 \\ 1 \end{bmatrix}$ for all $x_1, x_2 \in \mathbb{Z}_2$.

Solution.

Example 1.10.2

(b) $g : \mathbb{Z}_2^2 \to \mathbb{Z}_2^4$ given by $g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 \\ 1 \end{bmatrix}$ for all $x_1, x_2 \in \mathbb{Z}_2$.

Solution. (b) The function g is **not** linear because it does not satisfy axiom 1 of the definition of a linear function.

Example 1.10.2

(b) $g : \mathbb{Z}_2^2 \to \mathbb{Z}_2^4$ given by $g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 \\ 1 \end{bmatrix}$ for all $x_1, x_2 \in \mathbb{Z}_2$.

Solution. (b) The function g is **not** linear because it does not satisfy axiom 1 of the definition of a linear function. To see this, we consider, for example, the vectors $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in $\mathbb{Z}_2^2$, and we observe that (next slide):

Solution (continued).

$$g(\mathbf{u}+\mathbf{v}) \;=\; g\Big(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big) \;=\; g\Big(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\Big) \;=\; \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

whereas

$$g(\mathbf{u}) + g(\mathbf{v}) \;=\; g\Big(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big) + g\Big(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big) \;=\; \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$=\; \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

As we can see, $g(\mathbf{u}+\mathbf{v}) \neq g(\mathbf{u}) + g(\mathbf{v})$, and we deduce that g is not linear. □

Example 1.10.2

- (c) $h : \mathbb{Z}_3^3 \to \mathbb{Z}_3^2$ given by $h\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} x_1 + x_2 \\ x_1 x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{Z}_3$.

Solution.

Example 1.10.2

(c) $h : \mathbb{Z}_3^3 \to \mathbb{Z}_3^2$ given by $h\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{Z}_3$.

Solution. (c) The function h is **not** linear because it does not satisfy axiom 2 of the definition of a linear function.

Example 1.10.2

(c) $h : \mathbb{Z}_3^3 \to \mathbb{Z}_3^2$ given by $h\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{Z}_3$.

Solution. (c) The function h is **not** linear because it does not satisfy axiom 2 of the definition of a linear function. To see this, we consider, for example, the vector $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ in $\mathbb{Z}_3^3$ and the scalar $\alpha = 2$ in $\mathbb{Z}_3$, and we observe that

- $h(\alpha\mathbf{u}) = h\left(2\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right) = h\left(\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2+1 \\ 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$;
- $\alpha h(\mathbf{u}) = 2h\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right) = 2\begin{bmatrix} 1+2 \\ 1 \cdot 2 \end{bmatrix} = 2\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Thus, $h(\alpha\mathbf{u}) \neq \alpha h(\mathbf{u})$, and we deduce that h is not linear. $\square$

Proposition 1.10.3

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then $f(\mathbf{0}) = \mathbf{0}$.[a]

[a] Note that in $f(\mathbf{0}) = \mathbf{0}$, we have that $\mathbf{0} \in \mathbb{F}^m$, whereas $\mathbf{0} \in \mathbb{F}^n$.

Proof.

Proposition 1.10.3

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then $f(\mathbf{0}) = \mathbf{0}$.[a]

[a] Note that in $f(\mathbf{0}) = \mathbf{0}$, we have that $\mathbf{0} \in \mathbb{F}^m$, whereas $\mathbf{0} \in \mathbb{F}^n$.

Proof. We observe that

$$f(\mathbf{0}) \;=\; f(0 \cdot \mathbf{0}) \;\overset{(*)}{=}\; 0f(\mathbf{0}) \;=\; \mathbf{0},$$

where (*) follows from the fact that f is linear. □

Proposition 1.10.3

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then $f(\mathbf{0}) = \mathbf{0}$.[a]

[a] Note that in $f(\mathbf{0}) = \mathbf{0}$, we have that $\mathbf{0} \in \mathbb{F}^m$, whereas $\mathbf{0} \in \mathbb{F}^n$.

Proof. We observe that

$$f(\mathbf{0}) \;=\; f(0 \cdot \mathbf{0}) \;\overset{(*)}{=}\; 0f(\mathbf{0}) \;=\; \mathbf{0},$$

where (*) follows from the fact that f is linear. □

- **Remark:** Proposition 1.10.3 can sometimes be used to show that a function is not linear.

Proposition 1.10.3

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then $f(\mathbf{0}) = \mathbf{0}$.

Example 1.10.2

(b) $g : \mathbb{Z}_2^2 \to \mathbb{Z}_2^4$ given by $g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 \\ 1 \end{bmatrix}$ for all $x_1, x_2 \in \mathbb{Z}_2$.

- For example, for the function g from Example 1.10.2(b), we have that $g(\mathbf{0}) \neq \mathbf{0}$, and so g is not linear.

Proposition 1.10.3

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then $f(\mathbf{0}) = \mathbf{0}$.

Example 1.10.2

(b) $g : \mathbb{Z}_2^2 \to \mathbb{Z}_2^4$ given by $g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 \\ 1 \end{bmatrix}$ for all $x_1, x_2 \in \mathbb{Z}_2$.

- For example, for the function g from Example 1.10.2(b), we have that $g(\mathbf{0}) \neq \mathbf{0}$, and so g is not linear.
- However, note that the converse of Proposition 1.10.3: it is possible that a function $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is some field) satisfies $f(\mathbf{0}) = \mathbf{0}$, but that the function f is still not linear.

Proposition 1.10.3

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then $f(\mathbf{0}) = \mathbf{0}$.

Example 1.10.2

(c) $h : \mathbb{Z}_3^3 \to \mathbb{Z}_3^2$ given by $h\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 x_2 \end{bmatrix}$ for all $x_1, x_2, x_3 \in \mathbb{Z}_3$.

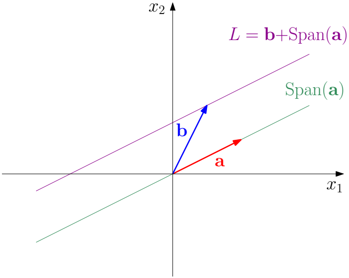
- For instance, the function h from Example 1.10.2(c) satisfies $h(\mathbf{0}) = \mathbf{0}$, but h is nevertheless not linear.

- Let us consider some geometric properties of linear functions from $\mathbb{R}^m$ to $\mathbb{R}^n$.

- Let us consider some geometric properties of linear functions from $\mathbb{R}^m$ to $\mathbb{R}^n$.
- Suppose that $f : \mathbb{R}^m \to \mathbb{R}^n$ is a linear function.

- Let us consider some geometric properties of linear functions from $\mathbb{R}^m$ to $\mathbb{R}^n$.
- Suppose that $f : \mathbb{R}^m \to \mathbb{R}^n$ is a linear function.
- It turns out that the image of any line in $\mathbb{R}^m$ under f is either a line in $\mathbb{R}^n$ or a point in $\mathbb{R}^n$ (technically, a set that contains only one point/vector of $\mathbb{R}^n$; we can think of such one-point sets as "degenerate lines").
    - This is one of the reasons why linear functions are called linear.

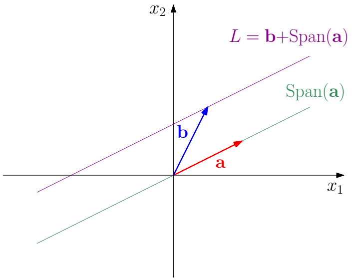
- Let us consider some geometric properties of linear functions from $\mathbb{R}^m$ to $\mathbb{R}^n$.
- Suppose that $f : \mathbb{R}^m \to \mathbb{R}^n$ is a linear function.
- It turns out that the image of any line in $\mathbb{R}^m$ under f is either a line in $\mathbb{R}^n$ or a point in $\mathbb{R}^n$ (technically, a set that contains only one point/vector of $\mathbb{R}^n$; we can think of such one-point sets as "degenerate lines").
    - This is one of the reasons why linear functions are called linear.
- Let us give a formal proof.



- Lines through the origin in $\mathbb{R}^m$ are simply sets of the form

$$\text{Span}(\mathbf{a}) \quad = \quad \{\alpha\mathbf{a} \mid \alpha \in \mathbb{R}\},$$

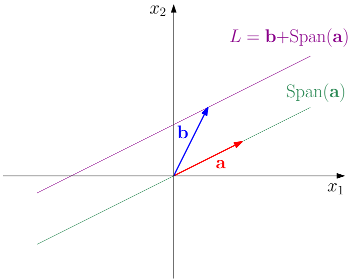
where **a** is a non-zero vector in $\mathbb{R}^m$.



- Lines through the origin in $\mathbb{R}^m$ are simply sets of the form

$$\text{Span}(\mathbf{a}) = \{\alpha\mathbf{a} \mid \alpha \in \mathbb{R}\},$$

  where $\mathbf{a}$ is a non-zero vector in $\mathbb{R}^m$.
- Any line in $\mathbb{R}^m$ is obtained by shifting a line through the origin by some vector $\mathbf{b} \in \mathbb{R}^m$ (if $\mathbf{b} = \mathbf{0}$, then our line still passes though the origin).



- Lines through the origin in $\mathbb{R}^m$ are simply sets of the form

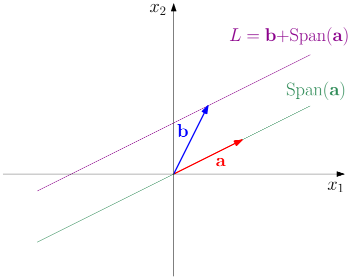
$$\text{Span}(\mathbf{a}) \quad = \quad \{\alpha \mathbf{a} \mid \alpha \in \mathbb{R}\},$$

where $\mathbf{a}$ is a non-zero vector in $\mathbb{R}^m$.

- Any line in $\mathbb{R}^m$ is obtained by shifting a line through the origin by some vector $\mathbf{b} \in \mathbb{R}^m$ (if $\mathbf{b} = \mathbf{0}$, then our line still passes though the origin).
- So, consider some line

$$L \quad := \quad \mathbf{b} + \text{Span}(\mathbf{a}) \quad = \quad \{\mathbf{b} + \alpha \mathbf{a} \mid \alpha \in \mathbb{R}\},$$

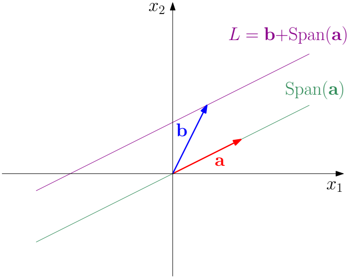
where $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{b}$ are fixed vectors in $\mathbb{R}^m$.



- For any point $\mathbf{b} + \alpha\mathbf{a}$ ($\alpha \in \mathbb{R}$) on the line L, we have that

$$f(\mathbf{b} + \alpha\mathbf{a}) \overset{(*)}{=} f(\mathbf{b}) + f(\alpha\mathbf{a}) \overset{(**)}{=} f(\mathbf{b}) + \alpha f(\mathbf{a})$$

where both (*) and (**) follow from the linearity of f, but in (*) we used axiom 1 from the definition of a linear function, and in (**) we used axiom 2.

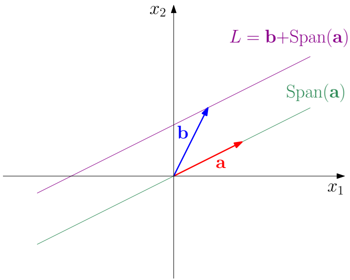


- For any point $\mathbf{b} + \alpha\mathbf{a}$ ($\alpha \in \mathbb{R}$) on the line L, we have that

$$f(\mathbf{b} + \alpha\mathbf{a}) \overset{(*)}{=} f(\mathbf{b}) + f(\alpha\mathbf{a}) \overset{(**)}{=} f(\mathbf{b}) + \alpha f(\mathbf{a})$$

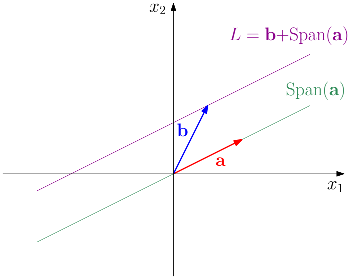
  where both (*) and (**) follow from the linearity of f, but in (*) we used axiom 1 from the definition of a linear function, and in (**) we used axiom 2.
- So, the image of our line L under f, denoted by $f[L]$, is

$$f[L] = \{f(\mathbf{b}) + \alpha f(\mathbf{a}) \mid \alpha \in \mathbb{R}\} = f(\mathbf{b}) + \text{Span}\big(f(\mathbf{a})\big).$$



- Reminder:

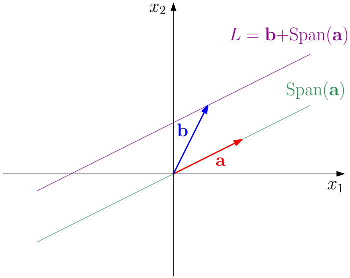
$$f[L] \;=\; \{f(\mathbf{b}) + \alpha f(\mathbf{a}) \mid \alpha \in \mathbb{R}\} \;=\; f(\mathbf{b}) + \mathrm{Span}\big(f(\mathbf{a})\big).$$



- Reminder:

$$f[L] \;=\; \{f(\mathbf{b}) + \alpha f(\mathbf{a}) \mid \alpha \in \mathbb{R}\} \;=\; f(\mathbf{b}) + \mathrm{Span}\big(f(\mathbf{a})\big).$$

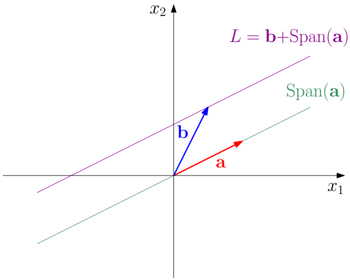
- If $f(\mathbf{a}) \neq \mathbf{0}$, then $f[L]$ is a line in $\mathbb{R}^n$.



- Reminder:

$$f[L] \quad = \quad \{f(\mathbf{b}) + \alpha f(\mathbf{a}) \mid \alpha \in \mathbb{R}\} \quad = \quad f(\mathbf{b}) + \operatorname{Span}\big(f(\mathbf{a})\big).$$

- If $f(\mathbf{a}) \neq \mathbf{0}$, then $f[L]$ is a line in $\mathbb{R}^n$.
- On the other hand, if $f(\mathbf{a}) = \mathbf{0}$, then $f[L] = \{f(\mathbf{b})\}$, which is a one-point subset ("degenerate line") of $\mathbb{R}^n$.



- Linear functions $f : \mathbb{R}^m \to \mathbb{R}^n$ map line segments onto line segments (possibly degenerate ones, i.e. those that contain only one point).
  - The proof is similar to the above and is left as an exercise.

- We note, however, that not all functions $f : \mathbb{R}^m \to \mathbb{R}^n$ that map lines to lines (or points) are linear.

- We note, however, that not all functions $f : \mathbb{R}^m \to \mathbb{R}^n$ that map lines to lines (or points) are linear.
- An obvious example might be a function $f : \mathbb{R}^n \to \mathbb{R}^n$ given by $f(\mathbf{x}) = \mathbf{x} + \mathbf{b}$ for all $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{b}$ is a fixed non-zero vector in $\mathbb{R}^n$.
    - This function is not linear because $f(\mathbf{0}) \neq \mathbf{0}$, and we know (by Proposition 1.10.3) that all linear functions map $\mathbf{0}$ to $\mathbf{0}$.

- We note, however, that not all functions $f : \mathbb{R}^m \to \mathbb{R}^n$ that map lines to lines (or points) are linear.
- An obvious example might be a function $f : \mathbb{R}^n \to \mathbb{R}^n$ given by $f(\mathbf{x}) = \mathbf{x} + \mathbf{b}$ for all $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{b}$ is a fixed non-zero vector in $\mathbb{R}^n$.
  - This function is not linear because $f(\mathbf{0}) \neq \mathbf{0}$, and we know (by Proposition 1.10.3) that all linear functions map $\mathbf{0}$ to $\mathbf{0}$.
- However, even if a function $f : \mathbb{R}^m \to \mathbb{R}^n$ maps lines to lines (or points) and maps $\mathbf{0}$ to $\mathbf{0}$, it might still fail to be linear.

- For example, consider the function $g : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^3 \\ 0 \end{bmatrix} \qquad \text{for all } x_1, x_2 \in \mathbb{R}.$$

- For example, consider the function $g : \mathbb{R}^2 \to \mathbb{R}^2$ given by

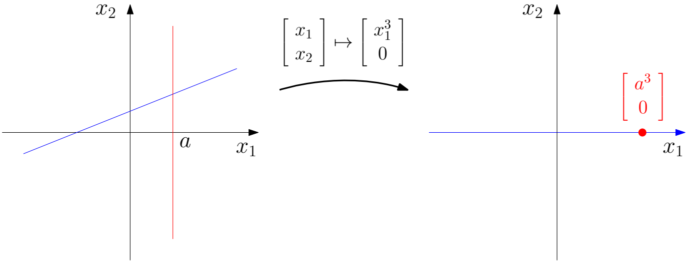
$$g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^3 \\ 0 \end{bmatrix} \qquad \text{for all } x_1, x_2 \in \mathbb{R}.$$

- This function is not linear, although it does map all lines onto either lines or points, and it does map $\mathbf{0}$ to $\mathbf{0}$.

- For example, consider the function $g : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^3 \\ 0 \end{bmatrix} \quad \text{for all } x_1, x_2 \in \mathbb{R}.$$

- This function is not linear, although it does map all lines onto either lines or points, and it does map $\mathbf{0}$ to $\mathbf{0}$.
- In particular, g maps any non-vertical line in $\mathbb{R}^2$ onto the x_1-axis, and it maps any vertical line onto a one-point set.



Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

Proof.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

Proof. By Corollary 1.7.6, the following hold:

(i) for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^m$, we have $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;

(ii) for all vectors $\mathbf{u} \in \mathbb{F}^m$ and scalars $\alpha \in \mathbb{F}$, we have that $A(\alpha\mathbf{u}) = \alpha(A\mathbf{u})$.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

Proof. By Corollary 1.7.6, the following hold:

(i) for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^m$, we have $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;

(ii) for all vectors $\mathbf{u} \in \mathbb{F}^m$ and scalars $\alpha \in \mathbb{F}$, we have that $A(\alpha\mathbf{u}) = \alpha(A\mathbf{u})$.

But now we have the following:

1. for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^m$, we have that

$$f(\mathbf{u} + \mathbf{v}) \;=\; A(\mathbf{u} + \mathbf{v}) \;\overset{(i)}{=}\; A\mathbf{u} + A\mathbf{v} \;=\; f(\mathbf{u}) + f(\mathbf{v});$$

2. for all vectors $\mathbf{u} \in \mathbb{F}^m$ and scalars $\alpha \in \mathbb{F}$, we have that

$$f(\alpha\mathbf{u}) \;=\; A(\alpha\mathbf{u}) \;\overset{(ii)}{=}\; \alpha(A\mathbf{u}) \;=\; \alpha f(\mathbf{u}).$$

So, f is linear. □

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- Mappings of the form $\mathbf{x} \mapsto A\mathbf{x}$, where A is some matrix, are sometimes called *matrix transformations*.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- Mappings of the form $\mathbf{x} \mapsto A\mathbf{x}$, where A is some matrix, are sometimes called *matrix transformations*.
- By Proposition 1.10.4, all matrix transformations are linear. Let us try to describe matrix transformations in a bit more detail.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- Mappings of the form $\mathbf{x} \mapsto A\mathbf{x}$, where A is some matrix, are sometimes called *matrix transformations*.
- By Proposition 1.10.4, all matrix transformations are linear. Let us try to describe matrix transformations in a bit more detail.
- Suppose we are given a matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times m}$ in $\mathbb{F}^{n\times m}$ (where $\mathbb{F}$ is some field), and define the function $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- Mappings of the form $\mathbf{x} \mapsto A\mathbf{x}$, where A is some matrix, are sometimes called *matrix transformations*.
- By Proposition 1.10.4, all matrix transformations are linear. Let us try to describe matrix transformations in a bit more detail.
- Suppose we are given a matrix $A = \left[\begin{array}{c} a_{i,j} \end{array} \right]_{n \times m}$ in $\mathbb{F}^{n \times m}$ (where $\mathbb{F}$ is some field), and define the function $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.
- But now for all vectors $\mathbf{x} = \left[\begin{array}{ccc} x_1 & \dots & x_m \end{array} \right]^T$ in $\mathbb{F}^m$, we have the following (next slide):

$$
\begin{aligned}
f(\mathbf{x}) &= A\mathbf{x} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \\
&= \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m \end{bmatrix}.
\end{aligned}
$$

$$
\begin{aligned}
f(\mathbf{x}) &= A\mathbf{x} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \\
&= \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m \end{bmatrix}.
\end{aligned}
$$

- So, our matrix transformation maps each vector $\mathbf{x} \in \mathbb{F}^m$ to a vector in $\mathbb{F}^n$, each of whose entries is a linear combination of the entries of $\mathbf{x}$, and the scalars/weights are determined by the corresponding row of the matrix A.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- By Proposition 1.10.4, every matrix transformation is a linear function.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- By Proposition 1.10.4, every matrix transformation is a linear function.
- Interestingly, a converse of sorts also holds.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- By Proposition 1.10.4, every matrix transformation is a linear function.
- Interestingly, a converse of sorts also holds.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix},$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proposition 1.10.4

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$ be a matrix, and define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a linear function.

- By Proposition 1.10.4, every matrix transformation is a linear function.
- Interestingly, a converse of sorts also holds.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix},$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

- We begin with another important theorem, which readily implies Theorem 1.10.6.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

- **Remark:** Theorem 1.10.5 essentially states that we can fully determine a linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field) by simply specifying what the standard basis vectors of $\mathbb{F}^m$ get mapped to.
  - Moreover, we can choose what the standard basis vectors get mapped to arbitrarily (i.e. we can map them to any vectors of $\mathbb{F}^n$ that we like).

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$.

Proof.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

Proof. **Existence.** Define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$.

Proof. **Existence.** Define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a matrix transformation, and so by Proposition 1.10.4, it is linear.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

Proof. **Existence.** Define $f : \mathbb{F}^m \to \mathbb{F}^n$ by setting $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then f is a matrix transformation, and so by Proposition 1.10.4, it is linear. Moreover, for all indices $i \in \{1, \dots, m\}$, we have that

$$f(\mathbf{e}_i) \;=\; A\mathbf{e}_i \;=\; \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \mathbf{e}_i \;\overset{(*)}{=}\; \mathbf{a}_i,$$

where (*) follows from Proposition 1.4.4.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

Proof (continued).

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$.

Proof (continued). **Uniqueness.** Suppose that $f : \mathbb{F}^m \to \mathbb{F}^n$ is **any** linear function that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$.

Proof (continued). **Uniqueness.** Suppose that $f : \mathbb{F}^m \to \mathbb{F}^n$ is **any** linear function that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$. WTS $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$.

Proof (continued). **Uniqueness.** Suppose that $f : \mathbb{F}^m \to \mathbb{F}^n$ is **any** linear function that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$. WTS $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Fix any vector $\mathbf{x} = \left[\begin{array}{ccc} x_1 & \dots & x_m \end{array} \right]^T$ in $\mathbb{F}^m$.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

Proof (continued). **Uniqueness.** Suppose that $f : \mathbb{F}^m \to \mathbb{F}^n$ is **any** linear function that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$. WTS $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Fix any vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_m \end{bmatrix}^T$ in $\mathbb{F}^m$. Then we have that

$$\mathbf{x} \quad = \quad x_1\mathbf{e}_1 + \dots + x_m\mathbf{e}_m,$$

and we compute (next slide):

Proof (continued). Reminder: $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$.

$$\begin{aligned}
f(\mathbf{x}) &= f\Big(x_1\mathbf{e}_1 + \dots + x_m\mathbf{e}_m\Big) \\
&\overset{(*)}{=} x_1 f(\mathbf{e}_1) + \dots + x_m f(\mathbf{e}_m) \\
&\overset{(**)}{=} x_1\mathbf{a}_1 + \dots + x_m\mathbf{a}_m \\
&\overset{(***)}{=} \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \\
&= A\mathbf{x},
\end{aligned}$$

where f follows from the linearity of f (and more precisely, from Proposition 1.10.1), (**) follows from the fact that $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, and (***) follows from the definition of matrix-vector multiplication. □

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$.

Theorem 1.10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{a}_1, \dots, \mathbf{a}_m$ be any vectors in $\mathbb{F}^n$. Then there exists a **unique** linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$, where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$. Moreover, this linear function f is given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$, where $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of* f) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix},$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof. **Existence.** Set $\mathbf{a}_1 := f(\mathbf{e}_1), \dots, \mathbf{a}_m := f(\mathbf{e}_m)$ and

$$A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix}.$$

Then by Theorem 1.10.5, we have that $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof. **Existence.** Set $\mathbf{a}_1 := f(\mathbf{e}_1), \dots, \mathbf{a}_m := f(\mathbf{e}_m)$ and

$$A := \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right] = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right].$$

Then by Theorem 1.10.5, we have that $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

- Indeed, $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix},$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof. **Existence.** Set $\mathbf{a}_1 := f(\mathbf{e}_1), \dots, \mathbf{a}_m := f(\mathbf{e}_m)$ and

$$A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix}.$$

Then by Theorem 1.10.5, we have that $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

- Indeed, $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$.
- So, by Theorem 1.10.5, we have that $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of* f) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof. **Existence.** Set $\mathbf{a}_1 := f(\mathbf{e}_1), \dots, \mathbf{a}_m := f(\mathbf{e}_m)$ and

$$A := \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right] = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right].$$

Then by Theorem 1.10.5, we have that $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

- Indeed, $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function that satisfies $f(\mathbf{e}_1) = \mathbf{a}_1, \dots, f(\mathbf{e}_m) = \mathbf{a}_m$.
- So, by Theorem 1.10.5, we have that $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

This proves existence.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A \quad = \quad \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof (continued). **Uniqueness.**

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof (continued). **Uniqueness.** Let $B = \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_m \end{array} \right]$ be **any** matrix in $\mathbb{F}^{n \times m}$ such that $f(\mathbf{x}) = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of f*) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \left[\begin{array}{ccc} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{array} \right],$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof (continued). **Uniqueness.** Let $B = \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_m \end{array} \right]$ be **any** matrix in $\mathbb{F}^{n \times m}$ such that $f(\mathbf{x}) = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then for all $i \in \{1, \dots, m\}$, we have that

$$f(\mathbf{e}_i) = B\mathbf{e}_i \overset{(*)}{=} \mathbf{b}_i$$

where (*) follows from Proposition 1.4.4.

Theorem 1.10.6

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function. Then there exists a unique matrix A (called the *standard matrix of* f) such that for all $\mathbf{x} \in \mathbb{F}^m$, we have that $f(\mathbf{x}) = A\mathbf{x}$. Moreover, the standard matrix A of f is given by

$$A = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix},$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors of $\mathbb{F}^m$.

Proof (continued). **Uniqueness.** Let $B = \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_m \end{bmatrix}$ be **any** matrix in $\mathbb{F}^{n \times m}$ such that $f(\mathbf{x}) = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^m$. Then for all $i \in \{1, \dots, m\}$, we have that

$$f(\mathbf{e}_i) = B\mathbf{e}_i \overset{(*)}{=} \mathbf{b}_i$$

where (*) follows from Proposition 1.4.4. Consequently,

$$B = \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_m \end{bmatrix} = \begin{bmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_m) \end{bmatrix}.$$

This proves uniqueness. □

Example 1.10.7

Find the standard matrix of the linear function $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$$

for all $x_1, x_2, x_3 \in \mathbb{R}$. (The fact that f is linear was proven in the solution of Example 1.10.2(a).)

Solution.

Example 1.10.7

Find the standard matrix of the linear function $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$$

for all $x_1, x_2, x_3 \in \mathbb{R}$. (The fact that f is linear was proven in the solution of Example 1.10.2(a).)

Solution. The standard matrix of f is

$$A := \begin{bmatrix} f(\mathbf{e}_1) & f(\mathbf{e}_2) & f(\mathbf{e}_3) \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

□

- For any set X, the *identity function* on X is the function $\mathrm{Id}_X : X \to X$ given by $\mathrm{Id}_X(x) = x$ for all $x \in X$.

- For any set X, the *identity function* on X is the function $\mathrm{Id}_X : X \to X$ given by $\mathrm{Id}_X(x) = x$ for all $x \in X$.

Proposition 1.10.8

Let $\mathbb{F}$ be a field. Then the identity function $\mathrm{Id}_{\mathbb{F}^n} : \mathbb{F}^n \to \mathbb{F}^n$ is linear, and its standard matrix is the identity matrix I_n.

Proof.

- For any set X, the *identity function* on X is the function $\mathrm{Id}_X : X \to X$ given by $\mathrm{Id}_X(x) = x$ for all $x \in X$.

Proposition 1.10.8

Let $\mathbb{F}$ be a field. Then the identity function $\mathrm{Id}_{\mathbb{F}^n} : \mathbb{F}^n \to \mathbb{F}^n$ is linear, and its standard matrix is the identity matrix I_n.

Proof. Obviously, the identity function $\mathrm{Id}_{\mathbb{F}^n}$ satisfies the two axioms from the definition of a linear function, and by Theorem 1.10.6, its standard matrix is

$$\left[\begin{array}{ccc} \mathrm{Id}_{\mathbb{F}^n}(\mathbf{e}_1) & \dots & \mathrm{Id}_{\mathbb{F}^n}(\mathbf{e}_n) \end{array}\right] = \left[\begin{array}{ccc} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{array}\right] = I_n.$$

- For any set X, the *identity function* on X is the function $\text{Id}_X : X \to X$ given by $\text{Id}_X(x) = x$ for all $x \in X$.

Proposition 1.10.8

Let $\mathbb{F}$ be a field. Then the identity function $\text{Id}_{\mathbb{F}^n} : \mathbb{F}^n \to \mathbb{F}^n$ is linear, and its standard matrix is the identity matrix I_n.

Proof. Obviously, the identity function $\text{Id}_{\mathbb{F}^n}$ satisfies the two axioms from the definition of a linear function, and by Theorem 1.10.6, its standard matrix is

$$\begin{bmatrix} \text{Id}_{\mathbb{F}^n}(\mathbf{e}_1) & \dots & \text{Id}_{\mathbb{F}^n}(\mathbf{e}_n) \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix} = I_n.$$

Alternatively, we observe that for any vector $\mathbf{x} \in \mathbb{F}^n$, we have that

$$\text{Id}_{\mathbb{F}^n}(\mathbf{x}) \stackrel{(*)}{=} \mathbf{x} \stackrel{(**)}{=} I_n\mathbf{x},$$

where (*) follows from the definition of the identity function, and (**) follows from Proposition 1.4.5.

- For any set X, the *identity function* on X is the function $\text{Id}_X : X \to X$ given by $\text{Id}_X(x) = x$ for all $x \in X$.

Proposition 1.10.8

Let $\mathbb{F}$ be a field. Then the identity function $\text{Id}_{\mathbb{F}^n} : \mathbb{F}^n \to \mathbb{F}^n$ is linear, and its standard matrix is the identity matrix I_n.

Proof. Obviously, the identity function $\text{Id}_{\mathbb{F}^n}$ satisfies the two axioms from the definition of a linear function, and by Theorem 1.10.6, its standard matrix is

$$\begin{bmatrix} \text{Id}_{\mathbb{F}^n}(\mathbf{e}_1) & \dots & \text{Id}_{\mathbb{F}^n}(\mathbf{e}_n) \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix} = I_n.$$

Alternatively, we observe that for any vector $\mathbf{x} \in \mathbb{F}^n$, we have that

$$\text{Id}_{\mathbb{F}^n}(\mathbf{x}) \overset{(*)}{=} \mathbf{x} \overset{(**)}{=} I_n\mathbf{x},$$

where (*) follows from the definition of the identity function, and (**) follows from Proposition 1.4.5. So, $\text{Id}_{\mathbb{F}^n}$ is a matrix transformation and is therefore linear (by Proposition 1.10.4), and its standard matrix is I_n. □

- Let us now consider some geometric examples (with pretty pictures).

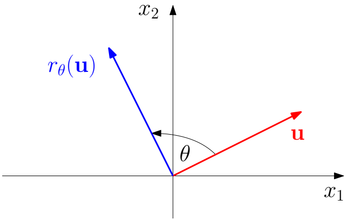
- Let us now consider some geometric examples (with pretty pictures).
- In particular, we consider a few special linear functions $f : \mathbb{R}^2 \to \mathbb{R}^2$.

- Let us now consider some geometric examples (with pretty pictures).
- In particular, we consider a few special linear functions $f : \mathbb{R}^2 \to \mathbb{R}^2$.
  - We will not formally prove that these functions are all linear.

- Let us now consider some geometric examples (with pretty pictures).
- In particular, we consider a few special linear functions $f : \mathbb{R}^2 \to \mathbb{R}^2$.
  - We will not formally prove that these functions are all linear.
  - To convince yourself that they are linear, think about what happens geometrically to sums and scalar multiples of vectors under these functions.

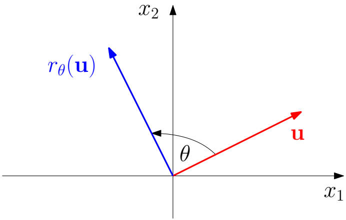
- **Rotation.** The function $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ that rotates each vector about the origin counterclockwise by the angle θ (see the picture below) is linear, and its standard matrix is

$$\begin{bmatrix} r_\theta(\mathbf{e}_1) & r_\theta(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$



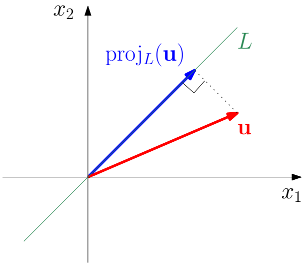
- **Rotation.** The function $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ that rotates each vector about the origin counterclockwise by the angle θ (see the picture below) is linear, and its standard matrix is

$$\left[\begin{array}{cc} r_\theta(\mathbf{e}_1) & r_\theta(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right].$$

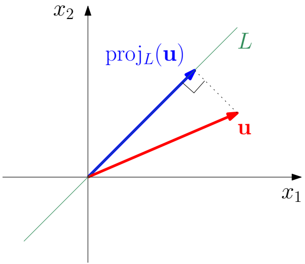


Note that rotating by the angle θ clockwise is the same as rotating by the angle $-\theta$ counterclockwise (which is why it is enough to consider only counterclockwise rotation, as long as we allow negative angles as well).

- **Orthogonal projection.** Given a line L in $\mathbb{R}^2$ that passes through the origin, the orthogonal projection $\text{proj}_L : \mathbb{R}^2 \to \mathbb{R}^2$ onto L (see the picture below) is linear.

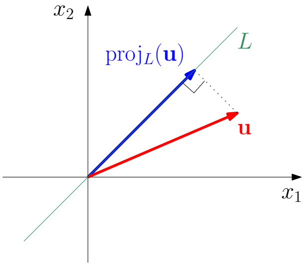


- **Orthogonal projection.** Given a line L in $\mathbb{R}^2$ that passes through the origin, the orthogonal projection $\text{proj}_L : \mathbb{R}^2 \to \mathbb{R}^2$ onto L (see the picture below) is linear.



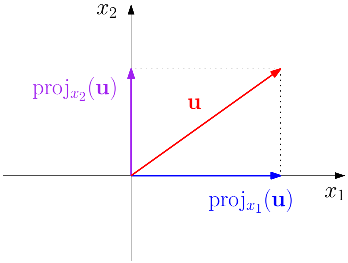
  - We cannot yet compute the standard matrix of orthogonal projection onto an arbitrary line through the origin; we will be able to do so only after we have developed a lot more theory.

- **Orthogonal projection.** Given a line L in $\mathbb{R}^2$ that passes through the origin, the orthogonal projection $\text{proj}_L : \mathbb{R}^2 \to \mathbb{R}^2$ onto L (see the picture below) is linear.

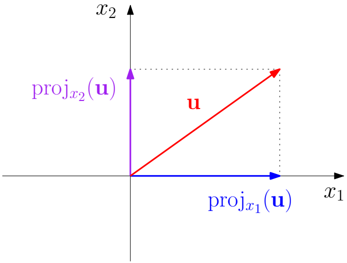


  - We cannot yet compute the standard matrix of orthogonal projection onto an arbitrary line through the origin; we will be able to do so only after we have developed a lot more theory.
  - However, we can already compute this matrix in some special cases.

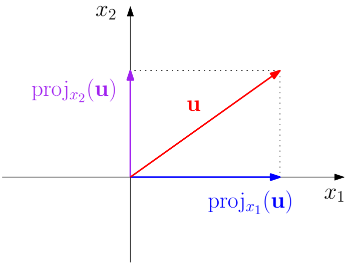
- **Orthogonal projection.**
  - Consider the projection $\text{proj}_{x_1} : \mathbb{R}^2 \to \mathbb{R}^2$ onto the x_1-axis and the projection $\text{proj}_{x_2} : \mathbb{R}^2 \to \mathbb{R}^2$ onto the x_2-axis.



- **Orthogonal projection.**
  - Consider the projection $\text{proj}_{x_1} : \mathbb{R}^2 \to \mathbb{R}^2$ onto the x_1-axis and the projection $\text{proj}_{x_2} : \mathbb{R}^2 \to \mathbb{R}^2$ onto the x_2-axis.



  - Note that for a vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ in $\mathbb{R}^2$, we have $\text{proj}_{x_1}(\mathbf{u}) = \begin{bmatrix} u_1 \\ 0 \end{bmatrix}$ and $\text{proj}_{x_2}(\mathbf{u}) = \begin{bmatrix} 0 \\ u_2 \end{bmatrix}$.



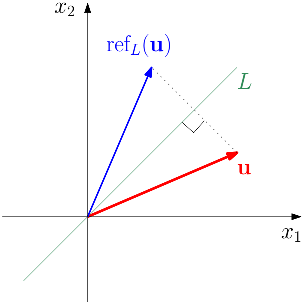
- The standard matrix of proj_{x_1} is

$$\begin{bmatrix} \text{proj}_{x_1}(\mathbf{e}_1) & \text{proj}_{x_1}(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

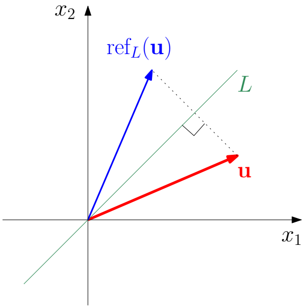
and the standard matrix of proj_{x_2} is

$$\begin{bmatrix} \text{proj}_{x_2}(\mathbf{e}_1) & \text{proj}_{x_2}(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

- **Reflection.** Given a line L in $\mathbb{R}^2$ that passes through the origin, the reflection $\text{ref}_L : \mathbb{R}^2 \to \mathbb{R}^2$ about the line L is linear.

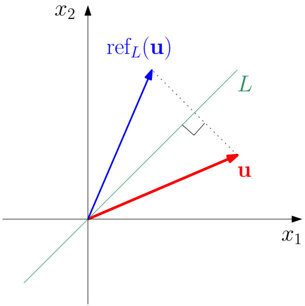


- **Reflection.** Given a line L in $\mathbb{R}^2$ that passes through the origin, the reflection $\text{ref}_L : \mathbb{R}^2 \to \mathbb{R}^2$ about the line L is linear.



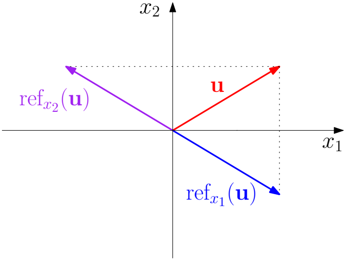
  - As in the case of orthogonal projections, we cannot yet compute the standard matrix of the reflection about an arbitrary line through the origin; we will only be able to do so once we have developed a lot more theory.

- **Reflection.** Given a line L in $\mathbb{R}^2$ that passes through the origin, the reflection $\mathrm{ref}_L : \mathbb{R}^2 \to \mathbb{R}^2$ about the line L is linear.

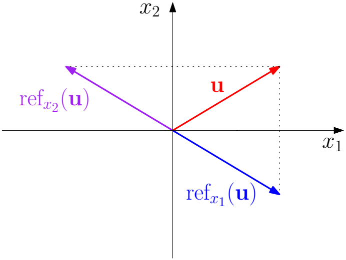


  - As in the case of orthogonal projections, we cannot yet compute the standard matrix of the reflection about an arbitrary line through the origin; we will only be able to do so once we have developed a lot more theory.
  - However, we can already compute this matrix in some special cases.

- **Reflection.**
  - Consider the reflection $\text{ref}_{x_1} : \mathbb{R}^2 \to \mathbb{R}^2$ about the x_1-axis and the reflection $\text{ref}_{x_2} : \mathbb{R}^2 \to \mathbb{R}^2$ about the x_2-axis.

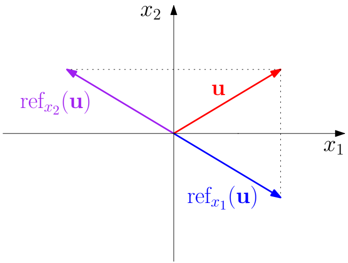


- **Reflection.**
  - Consider the reflection $\text{ref}_{x_1} : \mathbb{R}^2 \to \mathbb{R}^2$ about the x_1-axis and the reflection $\text{ref}_{x_2} : \mathbb{R}^2 \to \mathbb{R}^2$ about the x_2-axis.



  - Note that for a vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ in $\mathbb{R}^2$, we have

$$\text{ref}_{x_1}(\mathbf{u}) = \begin{bmatrix} u_1 \\ -u_2 \end{bmatrix} \text{ and } \text{ref}_{x_2}(\mathbf{u}) = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix}.$$



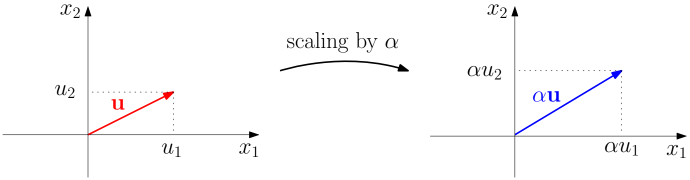
- The standard matrix of ref_{x_1} is

$$\left[\begin{array}{cc} \mathrm{ref}_{x_1}(\mathbf{e}_1) & \mathrm{ref}_{x_1}(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} \mathbf{e}_1 & -\mathbf{e}_2 \end{array}\right] = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right],$$

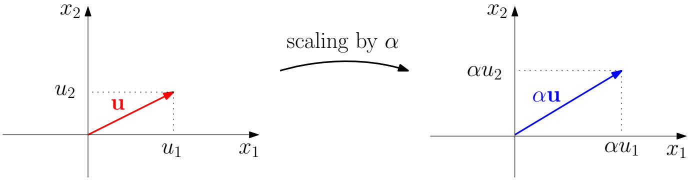
and the standard matrix of ref_{x_2} is

$$\left[\begin{array}{cc} \mathrm{ref}_{x_2}(\mathbf{e}_1) & \mathrm{ref}_{x_2}(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} -\mathbf{e}_1 & \mathbf{e}_2 \end{array}\right] = \left[\begin{array}{cc} -1 & 0 \\ 0 & 1 \end{array}\right].$$

- **Scaling.** Given a scalar $\alpha \in \mathbb{R}$, the function that scales each vector in $\mathbb{R}^2$ by α (see the picture below) is linear.



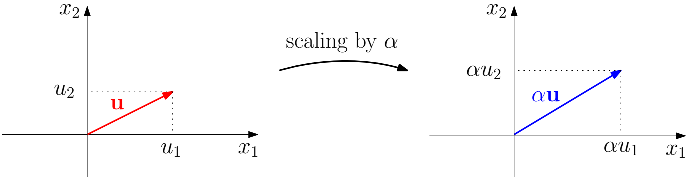
- **Scaling.** Given a scalar $\alpha \in \mathbb{R}$, the function that scales each vector in $\mathbb{R}^2$ by α (see the picture below) is linear.



  - The standard matrix of this linear function is

$$\begin{bmatrix} \alpha\mathbf{e}_1 & \alpha\mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}.$$

- **Scaling.** Given a scalar $\alpha \in \mathbb{R}$, the function that scales each vector in $\mathbb{R}^2$ by α (see the picture below) is linear.



  - The standard matrix of this linear function is

$$\begin{bmatrix} \alpha\mathbf{e}_1 & \alpha\mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}.$$

  - If $\alpha = 0$, then scaling by α is the same as mapping each vector to the origin.

- **Horizontal Shear.** A *horizontal shear* in $\mathbb{R}^2$ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by the formula

$$\mathbf{u} \mapsto \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \mathbf{u},$$

i.e. by the formula

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mapsto \begin{bmatrix} u_1 + ku_2 \\ u_2 \end{bmatrix},$$

where k is a fixed real constant.

- **Horizontal Shear.** A *horizontal shear* in $\mathbb{R}^2$ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by the formula

$$\mathbf{u} \mapsto \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \mathbf{u},$$

  i.e. by the formula

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mapsto \begin{bmatrix} u_1 + ku_2 \\ u_2 \end{bmatrix},$$

  where k is a fixed real constant.
    - This mapping has the effect of horizontally tilting objects in the coordinate plane (while keeping the vertical component unchanged).

- **Horizontal Shear.** A *horizontal shear* in $\mathbb{R}^2$ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by the formula

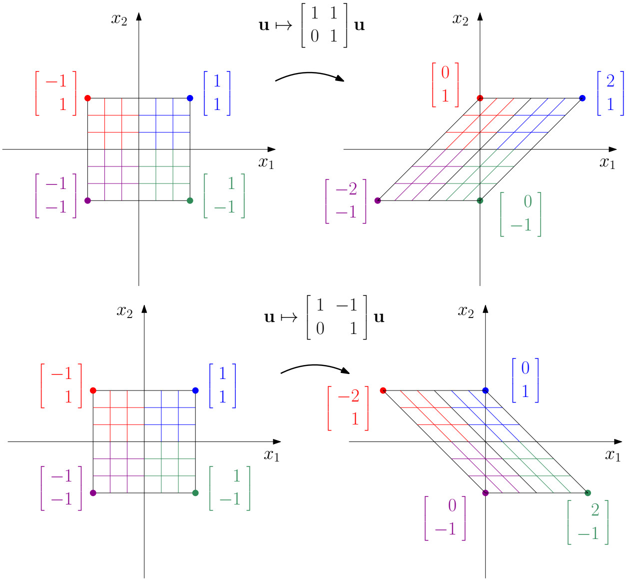
$$\mathbf{u} \mapsto \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \mathbf{u},$$

  i.e. by the formula

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mapsto \begin{bmatrix} u_1 + ku_2 \\ u_2 \end{bmatrix},$$

  where k is a fixed real constant.
  - This mapping has the effect of horizontally tilting objects in the coordinate plane (while keeping the vertical component unchanged).
  - This is illustrated on the next slide for the cases when $k = 1$ and $k = -1$.

- **Horizontal Shear.** $\mathbf{u} \mapsto \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \mathbf{u}$



- **Vertical Shear.** A *vertical shear* in $\mathbb{R}^2$ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by the formula

$$\mathbf{u} \mapsto \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \mathbf{u},$$

i.e. by the formula

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mapsto \begin{bmatrix} u_1 \\ ku_1 + u_2 \end{bmatrix},$$

where k is a fixed real constant.

- **Vertical Shear.** A *vertical shear* in $\mathbb{R}^2$ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by the formula

$$\mathbf{u} \mapsto \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \mathbf{u},$$

i.e. by the formula

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mapsto \begin{bmatrix} u_1 \\ ku_1 + u_2 \end{bmatrix},$$

where k is a fixed real constant.

  - This mapping has the effect of vertically tilting objects in the coordinate plane (while keeping the horizontal component unchanged).

- **Vertical Shear.** A *vertical shear* in $\mathbb{R}^2$ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by the formula

$$\mathbf{u} \mapsto \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \mathbf{u},$$

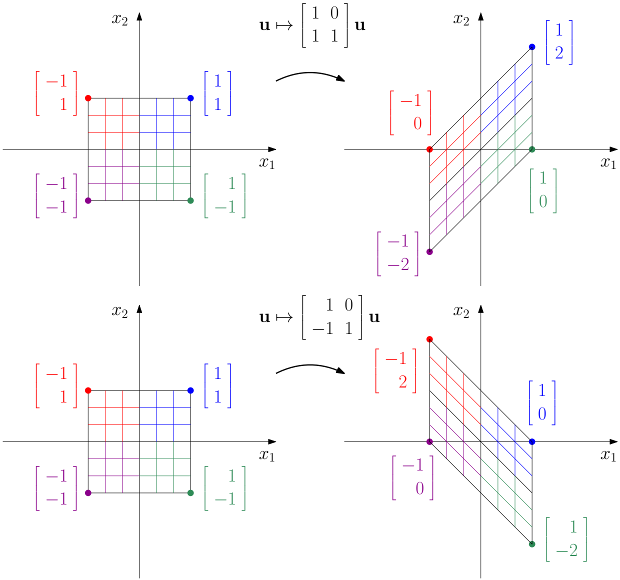
i.e. by the formula

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mapsto \begin{bmatrix} u_1 \\ ku_1 + u_2 \end{bmatrix},$$

where k is a fixed real constant.

  - This mapping has the effect of vertically tilting objects in the coordinate plane (while keeping the horizontal component unchanged).
  - This is illustrated on the next slide for the cases when $k = 1$ and $k = -1$.

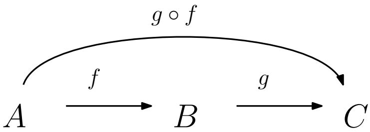
- **Vertical Shear.** $\mathbf{u} \mapsto \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \mathbf{u}$



- Given functions $f : A \to B$ and $g : B \to C$ (where A, B, and C are sets), we define the *composition of functions* g and f to be the function $g \circ f : A \to C$ given by

$$(g \circ f)(a) = g(f(a))$$

for all $a \in A$.



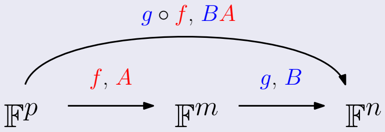
Proposition 1.10.1

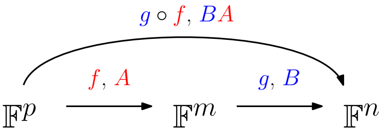
Let $\mathbb{F}$ be a field. Then all the following hold:

(a) for all linear functions $f, g : \mathbb{F}^m \to \mathbb{F}^n$, the function $f + g$ is linear, and moreover, if A and B (both in $\mathbb{F}^{n \times m}$) are the standard matrices of f and g, respectively, then $A + B$ is the standard matrix of $f + g$;

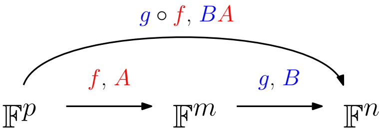
(b) for all linear functions $f : \mathbb{F}^m \to \mathbb{F}^n$ and scalars $\alpha \in \mathbb{F}$, the function αf is linear, and moreover, if $A \in \mathbb{F}^{n \times m}$ is the standard matrix of f, then αA is the standard matrix of αf;

(c) for all linear functions $f : \mathbb{F}^p \to \mathbb{F}^m$ and $g : \mathbb{F}^m \to \mathbb{F}^n$, the function $g \circ f$ is liner, and moreover, if $A \in \mathbb{F}^{m \times p}$ and $B \in \mathbb{F}^{n \times m}$ are the standard matrices of f and g, respectively, then BA is the standard matrix of $g \circ f$.

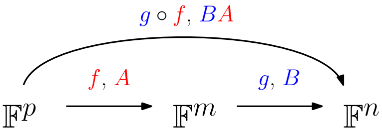




Proof of (c).



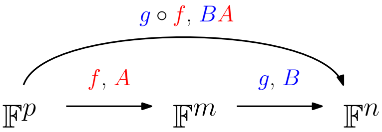
Proof of (c). Fix linear functions $f : \mathbb{F}^p \to \mathbb{F}^m$ and $g : \mathbb{F}^m \to \mathbb{F}^n$. Let $A \in \mathbb{F}^{m \times p}$ be the standard matrix of f, and let $B \in \mathbb{F}^{n \times m}$ be the standard matrix of g.



Proof of (c). Fix linear functions $f : \mathbb{F}^p \to \mathbb{F}^m$ and $g : \mathbb{F}^m \to \mathbb{F}^n$. Let $A \in \mathbb{F}^{m \times p}$ be the standard matrix of f, and let $B \in \mathbb{F}^{n \times m}$ be the standard matrix of g. Then for any $\mathbf{u} \in \mathbb{F}^p$, we have that

$$(g \circ f)(\mathbf{u}) \;=\; g(f(\mathbf{u})) \;\stackrel{(*)}{=}\; g(A\mathbf{u}) \;\stackrel{(**)}{=}\; B(A\mathbf{u}) \;\stackrel{(***)}{=}\; (BA)\mathbf{u},$$

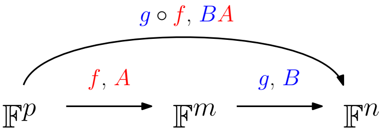
where (*) follows from the fact that A is the standard matrix of f, (**) follows from the fact that B is the standard matrix of g, and (***) follows from Corollary 1.7.6(g).



Proof of (c). Fix linear functions $f : \mathbb{F}^p \to \mathbb{F}^m$ and $g : \mathbb{F}^m \to \mathbb{F}^n$. Let $A \in \mathbb{F}^{m \times p}$ be the standard matrix of f, and let $B \in \mathbb{F}^{n \times m}$ be the standard matrix of g. Then for any $\mathbf{u} \in \mathbb{F}^p$, we have that

$$(g \circ f)(\mathbf{u}) = g(f(\mathbf{u})) \stackrel{(*)}{=} g(A\mathbf{u}) \stackrel{(**)}{=} B(A\mathbf{u}) \stackrel{(***)}{=} (BA)\mathbf{u},$$

where (*) follows from the fact that A is the standard matrix of f, (**) follows from the fact that B is the standard matrix of g, and (***) follows from Corollary 1.7.6(g). We have now shown that $g \circ f$ is a matrix transformation, and so (by Proposition 1.10.4) it is linear.



Proof of (c). Fix linear functions $f : \mathbb{F}^p \to \mathbb{F}^m$ and $g : \mathbb{F}^m \to \mathbb{F}^n$. Let $A \in \mathbb{F}^{m \times p}$ be the standard matrix of f, and let $B \in \mathbb{F}^{n \times m}$ be the standard matrix of g. Then for any $\mathbf{u} \in \mathbb{F}^p$, we have that

$$(g \circ f)(\mathbf{u}) \quad = \quad g(f(\mathbf{u})) \quad \overset{(*)}{=} \quad g(A\mathbf{u}) \quad \overset{(**)}{=} \quad B(A\mathbf{u}) \quad \overset{(***)}{=} \quad (BA)\mathbf{u},$$

where (*) follows from the fact that A is the standard matrix of f, (**) follows from the fact that B is the standard matrix of g, and (***) follows from Corollary 1.7.6(g). We have now shown that $g \circ f$ is a matrix transformation, and so (by Proposition 1.10.4) it is linear. Moreover, since (by the calculation above) we have that

$$(g \circ f)(\mathbf{u}) = (BA)\mathbf{u}$$

for all vectors $\mathbf{u} \in \mathbb{F}^p$, we see that BA is the standard matrix of $g \circ f$. □

Example 1.10.14

(a) Find the standard matrix of the linear function $f_1 : \mathbb{R}^2 \to \mathbb{R}^2$ that first reflects about the x_1-axis and then rotates about the origin counterclockwise by 90°.

(b) Find the standard matrix of the linear function $f_2 : \mathbb{R}^2 \to \mathbb{R}^2$ that first rotates about the origin counterclockwise by 90° and then reflects about the x_1-axis.

You may assume that f_1 and f_2 are indeed linear.

Solution.

Example 1.10.14

- (a) Find the standard matrix of the linear function $f_1 : \mathbb{R}^2 \to \mathbb{R}^2$ that first reflects about the x_1-axis and then rotates about the origin counterclockwise by 90°.
- (b) Find the standard matrix of the linear function $f_2 : \mathbb{R}^2 \to \mathbb{R}^2$ that first rotates about the origin counterclockwise by 90° and then reflects about the x_1-axis.

You may assume that f_1 and f_2 are indeed linear.

Solution. First, we note that the standard matrix of $\text{ref}_{x_1} : \mathbb{R}^2 \to \mathbb{R}^2$, the reflection about the x_1-axis, is

$$A = \begin{bmatrix} \text{ref}_{x_1}(\mathbf{e}_1) & \text{ref}_{x_1}(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & -\mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

whereas the standard matrix of $r_{90^\circ} : \mathbb{R}^2 \to \mathbb{R}^2$, the counterclockwise rotation by 90° about the origin, is

$$B = \begin{bmatrix} r_{90^\circ}(\mathbf{e}_1) & r_{90^\circ}(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \mathbf{e}_2 & -\mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Example 1.10.14

(a) Find the standard matrix of the linear function $f_1 : \mathbb{R}^2 \to \mathbb{R}^2$ that first reflects about the x_1-axis and then rotates about the origin counterclockwise by 90°.

(b) Find the standard matrix of the linear function $f_2 : \mathbb{R}^2 \to \mathbb{R}^2$ that first rotates about the origin counterclockwise by 90° and then reflects about the x_1-axis.

You may assume that f_1 and f_2 are indeed linear.

Solution (continued). Note that $f_1 = r_{90^\circ} \circ \text{ref}_{x_1}$ and $f_2 = \text{ref}_{x_1} \circ r_{90^\circ}$. So, by Proposition 1.10.13(c), the standard matrix of f_1 is

$$BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

whereas by the standard matrix of f_2 is

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

□

- We now briefly review one-to-one (injective) functions, onto (surjective) functions, and bijections.

- We now briefly review one-to-one (injective) functions, onto (surjective) functions, and bijections.
- This material properly belongs to Discrete Math.

- We now briefly review one-to-one (injective) functions, onto (surjective) functions, and bijections.
- This material properly belongs to Discrete Math.
- Here, we state (without proof) the results that we need.
  - However, all the proofs are in the Lecture Notes.

- We now briefly review one-to-one (injective) functions, onto (surjective) functions, and bijections.
- This material properly belongs to Discrete Math.
- Here, we state (without proof) the results that we need.
  - However, all the proofs are in the Lecture Notes.

Definition

A function $f : A \to B$ is said to be

- *one-to-one* (or *injective*, or an *injection*) if for all $a_1, a_2 \in A$ such that $a_1 \neq a_2$, we have $f(a_1) \neq f(a_2)$;[a]
- *onto* (or *surjective*, or a *surjection*) if for all $b \in B$, there exists some $a \in A$ such that $f(a) = b$;
- *bijective* or a *bijection* if it is both one-to-one and onto.

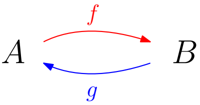
[a] Equivalently, $f : A \to B$ is *one-to-one* if for all $a_1, a_2 \in A$ such that $f(a_1) = f(a_2)$, we have that $a_1 = a_2$.

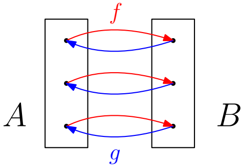
Proposition 1.10.15

Let $f : A \to B$ be a function. Then the following are equivalent:

(a) f is a bijection;

(b) there exists some function $g : B \to A$ such that $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$.



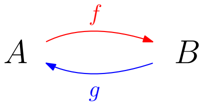


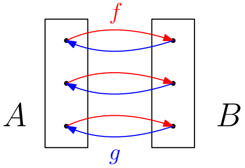
Proposition 1.10.15

Let $f : A \to B$ be a function. Then the following are equivalent:

(a) f is a bijection;

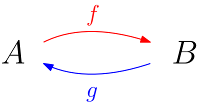
(b) there exists some function $g : B \to A$ such that $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$.

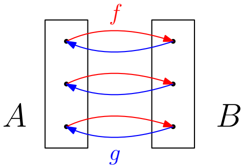




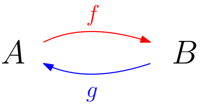
Proposition 1.10.16

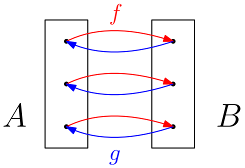
Let $f : A \to B$ be a bijection. Then there exists a **unique** function $g : B \to A$ such that $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$.



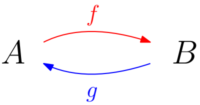


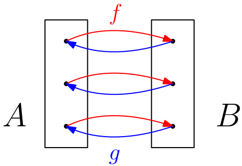
- **Terminology/Notation:** If $f : A \to B$ is a bijection, then the unique function $g : B \to A$ that satisfies $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$ (i.e. the function g from Proposition 1.10.16) is called the *inverse* of f and is denoted by f^{-1}.





- **Terminology/Notation:** If $f : A \to B$ is a bijection, then the unique function $g : B \to A$ that satisfies $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$ (i.e. the function g from Proposition 1.10.16) is called the *inverse* of f and is denoted by f^{-1}.
- Note that this means that:
  - $f^{-1} \circ f = \mathrm{Id}_A$;
  - $f \circ f^{-1} = \mathrm{Id}_B$;
  - for all $a \in A$ and $b \in B$, we have that $b = f(a)$ iff $a = f^{-1}(b)$.





- **Terminology/Notation:** If $f : A \to B$ is a bijection, then the unique function $g : B \to A$ that satisfies $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$ (i.e. the function g from Proposition 1.10.16) is called the *inverse* of f and is denoted by f^{-1}.
- Note that this means that:
  - $f^{-1} \circ f = \mathrm{Id}_A$;
  - $f \circ f^{-1} = \mathrm{Id}_B$;
  - for all $a \in A$ and $b \in B$, we have that $b = f(a)$ iff $a = f^{-1}(b)$.
- Note that the inverse of a bijection is also a bijection (by Proposition 1.10.15), and moreover, $(f^{-1})^{-1} = f$.

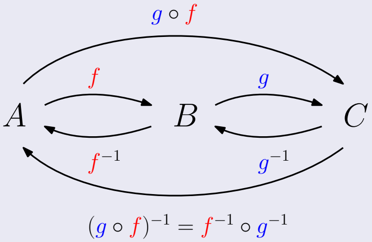
Proposition 1.10.17

Let $f : A \to B$ and $g : B \to C$ be functions. Then all the following hold:

a) if f and g are one-to-one, then $g \circ f$ is also one-to-one;

b) if f and g are onto, then $g \circ f$ is also onto;

c) if f and g are bijections, then $g \circ f$ is also a bijection, and moreover, $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ (see the diagram below).



- Back to linear functions!

- Back to linear functions!
- As the following theorem shows, we can easily check whether a linear function is one-to-one or onto by computing the rank of its standard matrix.

- Back to linear functions!
- As the following theorem shows, we can easily check whether a linear function is one-to-one or onto by computing the rank of its standard matrix.

Theorem 1.10.18

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then both the following hold:

(a) f is one-to-one iff $\text{rank}(A) = m$ (i.e. A has full column rank);

(b) f is onto iff $\text{rank}(A) = n$ (i.e. A has full row rank).

(a) f is one-to-one iff $\text{rank}(A) = m$ (i.e. A has full column rank)

Proof of (a).

(a) f is one-to-one iff $\text{rank}(A) = m$ (i.e. A has full column rank)

Proof of (a). We have the following sequence of equivalent statements:

$$
\begin{aligned}
f \text{ is one-to-one} &\overset{(*)}{\iff} \text{for all } \mathbf{b} \in \mathbb{F}^n,\ f(\mathbf{x}) = \mathbf{b} \text{ has at most one solution} \\
&\overset{(**)}{\iff} \text{for all } \mathbf{b} \in \mathbb{F}^n,\ A\mathbf{x} = \mathbf{b} \text{ has at most one solution,} \\
&\overset{(***)}{\iff} \text{rank}(A) = m,
\end{aligned}
$$

where (*) follows from the definition of a one-to-one function, (**) follows from the fact that A is the standard matrix of f, and (***) follows from Corollary 1.6.5. □

(b) f is onto iff $\text{rank}(A) = n$ (i.e. A has full row rank).

Proof of (b).

(b) f is onto iff $\text{rank}(A) = n$ (i.e. A has full row rank).

Proof of (b). (b) We have the following sequence of equivalent statements:

$$
\begin{aligned}
f \text{ is onto} \quad &\overset{(*)}{\iff} \quad \text{for all } \mathbf{b} \in \mathbb{F}^n,\ f(\mathbf{x}) = \mathbf{b} \text{ has at least one solution} \\
&\overset{(**)}{\iff} \quad \text{for all } \mathbf{b} \in \mathbb{F}^n,\ A\mathbf{x} = \mathbf{b} \text{ has at least one solution (i.e. } A\mathbf{x} = \mathbf{b} \text{ is consistent)} \\
&\overset{(***)}{\iff} \quad \text{rank}(A) = n,
\end{aligned}
$$

where (*) follows from the definition of an onto function, (**) follows from the fact that A is the standard matrix of f, and (***) follows from Corollary 1.6.6. □

Theorem 1.10.18

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then both the following hold:

- (a) f is one-to-one iff $\text{rank}(A) = m$ (i.e. A has full column rank);
- (b) f is onto iff $\text{rank}(A) = n$ (i.e. A has full row rank).

Definition

Let $\mathbb{F}$ be a field. A function $f : \mathbb{F}^m \to \mathbb{F}^n$ is an *isomorphism* if it is both linear and a bijection.

Definition

Let $\mathbb{F}$ be a field. A function $f : \mathbb{F}^m \to \mathbb{F}^n$ is an *isomorphism* if it is both linear and a bijection.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

Definition

Let $\mathbb{F}$ be a field. A function $f : \mathbb{F}^m \to \mathbb{F}^n$ is an *isomorphism* if it is both linear and a bijection.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

- By Theorem 1.10.19, if we know the standard matrix of a linear function, then we can easily determine whether that linear function is an isomorphism.

Definition

Let $\mathbb{F}$ be a field. A function $f : \mathbb{F}^m \to \mathbb{F}^n$ is an *isomorphism* if it is both linear and a bijection.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

- By Theorem 1.10.19, if we know the standard matrix of a linear function, then we can easily determine whether that linear function is an isomorphism.
- Moreover, Theorem 1.10.19 implies, in particular, that for a field $\mathbb{F}$, there can be no isomorphism from $\mathbb{F}^m$ to $\mathbb{F}^n$ for $m \neq n$.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

Proof.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

Proof. Suppose first that (a) holds. Since f is a one-to-one linear function, Theorem 1.10.18(a) guarantees that $\text{rank}(A) = m$.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

Proof. Suppose first that (a) holds. Since f is a one-to-one linear function, Theorem 1.10.18(a) guarantees that $\text{rank}(A) = m$. On the other hand, since f is an onto linear function, Theorem 1.10.18(b) guarantees that $\text{rank}(A) = n$.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

Proof. Suppose first that (a) holds. Since f is a one-to-one linear function, Theorem 1.10.18(a) guarantees that $\text{rank}(A) = m$. On the other hand, since f is an onto linear function, Theorem 1.10.18(b) guarantees that $\text{rank}(A) = n$. But now $m = \text{rank}(A) = n$, and (b) follows.

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

Proof. Suppose first that (a) holds. Since f is a one-to-one linear function, Theorem 1.10.18(a) guarantees that $\text{rank}(A) = m$. On the other hand, since f is an onto linear function, Theorem 1.10.18(b) guarantees that $\text{rank}(A) = n$. But now $m = \text{rank}(A) = n$, and (b) follows.

Suppose now that (b) holds. Then by Theorem 1.10.18(a), f is one-to-one, and by Theorem 1.10.18(b), f is onto. So, f is a bijection. Since f is also linear (by hypothesis), we deduce that f is an isomorphism, i.e. (a) holds. $\square$

Theorem 1.10.18

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then both the following hold:

(a) f is one-to-one iff $\text{rank}(A) = m$ (i.e. A has full column rank);

(b) f is onto iff $\text{rank}(A) = n$ (i.e. A has full row rank).

Theorem 1.10.19

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^m \to \mathbb{F}^n$ be a linear function, and let $A \in \mathbb{F}^{n \times m}$ be the standard matrix of f. Then the following are equivalent:

(a) f is an isomorphism;

(b) $\text{rank}(A) = m = n$ (i.e. A is a square matrix of full rank).

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof.

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof. Since $f : \mathbb{F}^n \to \mathbb{F}^n$ is an isomorphism, it is, in particular, a bijection; consequently, f has an inverse $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$, which is also a bijection.

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof. Since $f : \mathbb{F}^n \to \mathbb{F}^n$ is an isomorphism, it is, in particular, a bijection; consequently, f has an inverse $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$, which is also a bijection. So, to show that f^{-1} is an isomorphism, it suffices to show that f^{-1} is linear.

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof. Since $f : \mathbb{F}^n \to \mathbb{F}^n$ is an isomorphism, it is, in particular, a bijection; consequently, f has an inverse $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$, which is also a bijection. So, to show that f^{-1} is an isomorphism, it suffices to show that f^{-1} is linear.

First, fix $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{F}^n$. WTS $f^{-1}(\mathbf{v}_1 + \mathbf{v}_2) = f^{-1}(\mathbf{v}_1) + f^{-1}(\mathbf{v}_2)$.

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof. Since $f : \mathbb{F}^n \to \mathbb{F}^n$ is an isomorphism, it is, in particular, a bijection; consequently, f has an inverse $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$, which is also a bijection. So, to show that f^{-1} is an isomorphism, it suffices to show that f^{-1} is linear.

First, fix $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{F}^n$. WTS $f^{-1}(\mathbf{v}_1 + \mathbf{v}_2) = f^{-1}(\mathbf{v}_1) + f^{-1}(\mathbf{v}_2)$. Set $\mathbf{u}_1 := f^{-1}(\mathbf{v}_1)$ and $\mathbf{u}_2 := f^{-1}(\mathbf{v}_2)$, so that $f(\mathbf{u}_1) = \mathbf{v}_1$ and $f(\mathbf{u}_2) = \mathbf{v}_2$.

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof. Since $f : \mathbb{F}^n \to \mathbb{F}^n$ is an isomorphism, it is, in particular, a bijection; consequently, f has an inverse $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$, which is also a bijection. So, to show that f^{-1} is an isomorphism, it suffices to show that f^{-1} is linear.

First, fix $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{F}^n$. WTS $f^{-1}(\mathbf{v}_1 + \mathbf{v}_2) = f^{-1}(\mathbf{v}_1) + f^{-1}(\mathbf{v}_2)$. Set $\mathbf{u}_1 := f^{-1}(\mathbf{v}_1)$ and $\mathbf{u}_2 := f^{-1}(\mathbf{v}_2)$, so that $f(\mathbf{u}_1) = \mathbf{v}_1$ and $f(\mathbf{u}_2) = \mathbf{v}_2$. Then (next slide):

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof (continued). Reminder: $\mathbf{u}_1 := f^{-1}(\mathbf{v}_1)$ and $\mathbf{u}_2 := f^{-1}(\mathbf{v}_2)$, so that $f(\mathbf{u}_1) = \mathbf{v}_1$ and $f(\mathbf{u}_2) = \mathbf{v}_2$.

$$\begin{aligned}
f^{-1}(\mathbf{v}_1 + \mathbf{v}_2) &= f^{-1}\big(f(\mathbf{u}_1) + f(\mathbf{u}_2)\big) \\
&= f^{-1}\big(f(\mathbf{u}_1 + \mathbf{u}_2)\big) && \text{because } f \text{ is linear} \\
&= (f^{-1} \circ f)(\mathbf{u}_1 + \mathbf{u}_2) \\
&= \mathrm{Id}_{\mathbb{F}^n}(\mathbf{u}_1 + \mathbf{u}_2) \\
&= \mathbf{u}_1 + \mathbf{u}_2 \\
&= f^{-1}(\mathbf{v}_1) + f^{-1}(\mathbf{v}_2).
\end{aligned}$$

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof (continued). Next, fix $\mathbf{v} \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$. WTS $f^{-1}(\alpha\mathbf{v}) = \alpha f^{-1}(\mathbf{v})$.

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof (continued). Next, fix $\mathbf{v} \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$. WTS $f^{-1}(\alpha\mathbf{v}) = \alpha f^{-1}(\mathbf{v})$. Set $\mathbf{u} := f^{-1}(\mathbf{v})$, so that $f(\mathbf{u}) = \mathbf{v}$.

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof (continued). Next, fix $\mathbf{v} \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$. WTS $f^{-1}(\alpha\mathbf{v}) = \alpha f^{-1}(\mathbf{v})$. Set $\mathbf{u} := f^{-1}(\mathbf{v})$, so that $f(\mathbf{u}) = \mathbf{v}$. Then

$$\begin{aligned}
f^{-1}(\alpha\mathbf{v}) &= f^{-1}\big(\alpha f(\mathbf{u})\big) \\
&= f^{-1}\big(f(\alpha\mathbf{u})\big) && \text{because } f \text{ is linear} \\
&= (f^{-1} \circ f)(\alpha\mathbf{u}) \\
&= \mathrm{Id}_{\mathbb{F}^n}(\alpha\mathbf{u}) \\
&= \alpha\mathbf{u} \\
&= \alpha f^{-1}(\mathbf{v}).
\end{aligned}$$

Theorem 1.10.20

Let $\mathbb{F}$ be a field, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be an isomorphism. Then $f^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ is also an isomorphism.

Proof (continued). Next, fix $\mathbf{v} \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$. WTS $f^{-1}(\alpha\mathbf{v}) = \alpha f^{-1}(\mathbf{v})$. Set $\mathbf{u} := f^{-1}(\mathbf{v})$, so that $f(\mathbf{u}) = \mathbf{v}$. Then

$$
\begin{aligned}
f^{-1}(\alpha\mathbf{v}) &= f^{-1}\big(\alpha f(\mathbf{u})\big) \\
&= f^{-1}\big(f(\alpha\mathbf{u})\big) && \text{because } f \text{ is linear} \\
&= (f^{-1} \circ f)(\alpha\mathbf{u}) \\
&= \mathrm{Id}_{\mathbb{F}^n}(\alpha\mathbf{u}) \\
&= \alpha\mathbf{u} \\
&= \alpha f^{-1}(\mathbf{v}).
\end{aligned}
$$

We have now proven that f^{-1} linear. This completes the argument. $\square$