Linear Algebra 2

Lecture #12

Matrices of linear functions between non-trivial, finite-dimensional vector spaces

Irena Penev

February 18-19, 2025

- This lecture has four parts:

- This lecture has four parts:
  1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces

- This lecture has four parts:
  1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces
  2. Change of basis (transition) matrices

- This lecture has four parts:
  1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces
  2. Change of basis (transition) matrices
  3. Similar matrices

- This lecture has four parts:
  1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces
  2. Change of basis (transition) matrices
  3. Similar matrices
  4. Checking the existence and uniqueness of linear functions with certain specifications: examples with polynomials and matrices

1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces

1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces

Theorem 1.10.5

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*) s.t. 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 $\mathbb{F}^m$.

1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces

Theorem 1.10.5

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) s.t. 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 $\mathbb{F}^m$.

- Linear functions between general vector spaces do **not** have standard matrices.

1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces

Theorem 1.10.5

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*) s.t. 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 $\mathbb{F}^m$.

- Linear functions between general vector spaces do **not** have standard matrices.
- However, we can associate a matrix to a linear function between non-trivial, finite-dimensional vector spaces, provided we have first specified a basis of the domain and a basis of the codomain.

1. Matrices of linear functions between non-trivial, finite-dimensional vector spaces

Theorem 1.10.5

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) s.t. 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 $\mathbb{F}^m$.

- Linear functions between general vector spaces do **not** have standard matrices.
- However, we can associate a matrix to a linear function between non-trivial, finite-dimensional vector spaces, provided we have first specified a basis of the domain and a basis of the codomain.
- First, we review some of the results from previous lectures.

Theorem 3.2.7

Let V be a vector space over a field $\mathbb{F}$, and let $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$. Then the following are equivalent:

(i) $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of V;

(ii) for all vectors $\mathbf{v} \in V$, there exist **unique** scalars $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ s.t. $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n$.

Theorem 3.2.7

Let V be a vector space over a field $\mathbb{F}$, and let $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$. Then the following are equivalent:

(i) $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of V;

(ii) for all vectors $\mathbf{v} \in V$, there exist **unique** scalars $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ s.t. $\mathbf{v} = \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n$.

- Suppose $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ $(n \geq 1)$ is a basis of a vector space V over a field $\mathbb{F}$. Then by Theorem 3.2.7, to every vector $\mathbf{v} \in V$, we can associate a unique vector

$$\left[\ \mathbf{v}\ \right]_{\mathcal{B}} := \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

in $\mathbb{F}^n$ s.t. $\mathbf{v} = \alpha_1\mathbf{b}_1 + \dots + \alpha_n\mathbf{b}_n$; the vector $\left[\ \mathbf{v}\ \right]_{\mathcal{B}}$ is called the *coordinate vector* of $\mathbf{v}$ associated with the basis $\mathcal{B}$.

- **Remark:** Suppose that $\mathbb{F}$ is a field, and that $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{F}^n$.
  - Then for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that

$$\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n,$$

and consequently,

$$\begin{bmatrix} \mathbf{x} \end{bmatrix}_{\mathcal{E}_n} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T = \mathbf{x}.$$

Proposition 3.2.9

Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ $(n \geq 1)$ be a basis of a vector space V over a field $\mathbb{F}$. Then for all $i \in \{1, \dots, n\}$, we have that $\left[\ \mathbf{b}_i\ \right]_{\mathcal{B}} = \mathbf{e}_i^n$.

Proof. Fix $i \in \{1, \dots, n\}$. Then

$$\mathbf{b}_i \quad = \quad 0\mathbf{b}_1 + \dots + 0\mathbf{b}_{i-1} + 1\mathbf{b}_i + 0\mathbf{b}_{i+1} + \dots + 0\mathbf{b}_n$$

and consequently,

$$\left[\ \mathbf{b}_i\ \right]_{\mathcal{B}} \quad = \quad \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{-th entry}$$

i.e. $\left[\ \mathbf{b}_i\ \right]_{\mathcal{B}} = \mathbf{e}_i^n$. □

Theorem 4.3.2

Let U and V be vector spaces over a field $\mathbb{F}$, and assume that U is finite-dimensional. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be a basis of U, and let $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$.[a] Then there exists a unique linear function $f : U \to V$ s.t. $f(\mathbf{u}_1) = \mathbf{v}_1, \dots, f(\mathbf{u}_n) = \mathbf{v}_n$. Moreover, if the vector space U is non-trivial (i.e. $n \neq 0$), then this unique linear function $f : U \to V$ satisfies the following: for all $\mathbf{u} \in U$, we have that

$$f(\mathbf{u}) \;=\; \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n,$$

where $\left[\ \mathbf{u}\ \right]_{\mathcal{B}} = \left[\ \alpha_1 \quad \dots \quad \alpha_n\ \right]^T$. On the other hand, if U is trivial (i.e. $U = \{\mathbf{0}\}$),[b] then $f : U \to V$ is given by $f(\mathbf{0}) = \mathbf{0}$.

---

[a] Here, $\mathbf{v}_1, \dots, \mathbf{v}_n$ are arbitrary vectors in V. They are not necessarily pairwise distinct.

[b] Note that in this case, we have that $n = 0$ and $\mathcal{B} = \emptyset$.

Corollary 4.3.3

Let U and V be vector spaces over a field $\mathbb{F}$, and assume that U is finite-dimensional. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a linearly independent set of vectors in U, and let $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$.[a] Then there exists a linear function $f : U \to V$ s.t. $f(\mathbf{u}_1) = \mathbf{v}_1, \dots, f(\mathbf{u}_k) = \mathbf{v}_k$. Moreover, if V is non-trivial, then this linear function f is unique iff $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of U.

---

[a] Here, $\mathbf{v}_1, \dots, \mathbf{v}_k$ are arbitrary vectors in V. They are not necessarily pairwise distinct.

- Standard matrices once again:

Theorem 1.10.5

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*) s.t. 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 $\mathbb{F}^m$.

- Let's generalize this (next slide)!

Theorem 4.5.1

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function. Then exists a unique matrix in $\mathbb{F}^{n \times m}$, denoted by ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$ and called the *matrix of f with respect to $\mathcal{B}$ and $\mathcal{C}$*, s.t. for all $\mathbf{u} \in U$, we have that

$$ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} \ [\ \mathbf{u}\ ]_{\mathcal{B}} \ = \ [\ f(\mathbf{u})\ ]_{\mathcal{C}}. $$

Moreover, the matrix ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$ is given by

$$ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} \ = \ \left[\ [\ f(\mathbf{b}_1)\ ]_{\mathcal{C}} \ \dots \ [\ f(\mathbf{b}_m)\ ]_{\mathcal{C}}\ \right]. $$

Theorem 4.5.1

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function. Then exists a unique matrix in $\mathbb{F}^{n \times m}$, denoted by ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ and called the *matrix of f with respect to $\mathcal{B}$ and $\mathcal{C}$*, s.t. for all $\mathbf{u} \in U$, we have that

$$ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\ \big[\ \mathbf{u}\ \big]_{\mathcal{B}} \ = \ \big[\ f(\mathbf{u})\ \big]_{\mathcal{C}}. $$

Moreover, the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is given by

$$ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \ = \ \Big[\ \big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}} \ \ \dots \ \ \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\ \Big]. $$

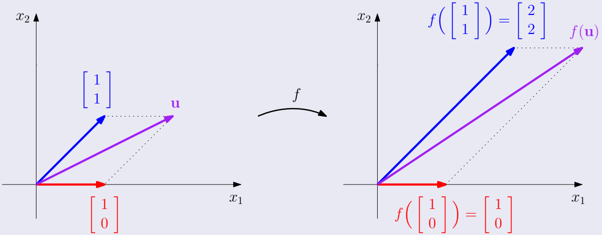
- First an example and a remark, then a proof.

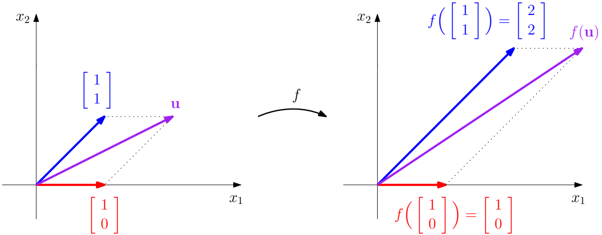
Example 4.5.2

Consider the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ of $\mathbb{R}^2$, and consider the unique linear function $f : \mathbb{R}^2 \to \mathbb{R}^2$ that satisfies the following:

- $f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$,
- $f\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

Compute the matrix ${}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}}$.





- **Remark:** The fact that $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ is a basis of $\mathbb{R}^2$ follows from the fact that $\operatorname{rank}\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right) = 2$ and from the Invertible Matrix Theorem. The existence and uniqueness of the linear function f follows from Theorem 4.3.2.

Solution. Using the formula from Theorem 4.5.1, we compute:

$$
\begin{aligned}
{}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}} &= \left[\ \left[\ f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)\ \right]_{\mathcal{B}} \quad \left[\ f\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)\ \right]_{\mathcal{B}}\ \right] \\
&= \left[\ \left[\ \begin{bmatrix} 1 \\ 0 \end{bmatrix}\ \right]_{\mathcal{B}} \quad \left[\ \begin{bmatrix} 2 \\ 2 \end{bmatrix}\ \right]_{\mathcal{B}}\ \right] \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.
\end{aligned}
$$

□

- **Remark:** Matrices of the form ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$ are generalizations of standard matrices.

- **Remark:** Matrices of the form ${}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$ are generalizations of standard matrices.
    - Indeed, if $\mathbb{F}$ is a field and $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function, then the matrix

$$ {}_{\mathcal{E}_n}\left[\ f\ \right]_{\mathcal{E}_m} $$

      is precisely the **standard matrix** of f, where as usual, $\mathcal{E}_m = \{\mathbf{e}_1^m, \dots, \mathbf{e}_m^m\}$ is the standard basis of $\mathbb{F}^m$, and $\mathcal{E}_n = \{\mathbf{e}_1^n, \dots, \mathbf{e}_n^n\}$ is the standard basis of $\mathbb{F}^n$.

- **Remark:** Matrices of the form ${}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$ are generalizations of standard matrices.
    - Indeed, if $\mathbb{F}$ is a field and $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function, then the matrix

$$ {}_{\mathcal{E}_n}\left[\ f\ \right]_{\mathcal{E}_m} $$

      is precisely the **standard matrix** of f, where as usual, $\mathcal{E}_m = \{\mathbf{e}_1^m, \dots, \mathbf{e}_m^m\}$ is the standard basis of $\mathbb{F}^m$, and $\mathcal{E}_n = \{\mathbf{e}_1^n, \dots, \mathbf{e}_n^n\}$ is the standard basis of $\mathbb{F}^n$.
    - Indeed, suppose that $\mathbb{F}$ be a field, that $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function, and that A is the standard matrix of f.

- **Remark:** Matrices of the form ${}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$ are generalizations of standard matrices.
  - Indeed, if $\mathbb{F}$ is a field and $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function, then the matrix
$$ {}_{\mathcal{E}_n}\left[\ f\ \right]_{\mathcal{E}_m} $$
is precisely the **standard matrix** of f, where as usual, $\mathcal{E}_m = \{\mathbf{e}_1^m, \dots, \mathbf{e}_m^m\}$ is the standard basis of $\mathbb{F}^m$, and $\mathcal{E}_n = \{\mathbf{e}_1^n, \dots, \mathbf{e}_n^n\}$ is the standard basis of $\mathbb{F}^n$.
  - Indeed, suppose that $\mathbb{F}$ be a field, that $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function, and that A is the standard matrix of f.
  - Then for all $\mathbf{u} \in \mathbb{F}^m$, we have the following:
$$ A\left[\ \mathbf{u}\ \right]_{\mathcal{E}_m} \;=\; A\mathbf{u} \;=\; f(\mathbf{u}) \;=\; \left[\ f(\mathbf{u})\ \right]_{\mathcal{E}_n}. $$

- **Remark:** Matrices of the form ${}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$ are generalizations of standard matrices.
  - Indeed, if $\mathbb{F}$ is a field and $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function, then the matrix

    $${}_{\mathcal{E}_n}\left[\ f\ \right]_{\mathcal{E}_m}$$

    is precisely the **standard matrix** of f, where as usual, $\mathcal{E}_m = \{\mathbf{e}_1^m, \dots, \mathbf{e}_m^m\}$ is the standard basis of $\mathbb{F}^m$, and $\mathcal{E}_n = \{\mathbf{e}_1^n, \dots, \mathbf{e}_n^n\}$ is the standard basis of $\mathbb{F}^n$.
  - Indeed, suppose that $\mathbb{F}$ be a field, that $f : \mathbb{F}^m \to \mathbb{F}^n$ is a linear function, and that A is the standard matrix of f.
  - Then for all $\mathbf{u} \in \mathbb{F}^m$, we have the following:

    $$A\left[\ \mathbf{u}\ \right]_{\mathcal{E}_m} \quad = \quad A\mathbf{u} \quad = \quad f(\mathbf{u}) \quad = \quad \left[\ f(\mathbf{u})\ \right]_{\mathcal{E}_n}.$$

  - Now the uniqueness part of Theorem 4.5.1 guarantees that $A = {}_{\mathcal{E}_n}\left[\ f\ \right]_{\mathcal{E}_m}$, i.e. ${}_{\mathcal{E}_n}\left[\ f\ \right]_{\mathcal{E}_m}$ is the standard matrix of f.

Theorem 4.5.1

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function. Then exists a unique matrix in $\mathbb{F}^{n\times m}$, denoted by ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ and called the *matrix of f with respect to $\mathcal{B}$ and $\mathcal{C}$*, s.t. for all $\mathbf{u} \in U$, we have that

$$ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \ \big[\ \mathbf{u}\ \big]_{\mathcal{B}} \ = \ \big[\ f(\mathbf{u})\ \big]_{\mathcal{C}}. $$

Moreover, the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is given by

$$ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \ = \ \Big[\ \big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}} \ \dots \ \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\ \Big]. $$

- Let's prove the theorem!

Proof. **Existence.**

Proof. **Existence.** Fix $\mathbf{u} \in U$. We must show that

$$\left[\left[f(\mathbf{b}_1) \right]_{\mathcal{C}} \quad \ldots \quad \left[f(\mathbf{b}_m) \right]_{\mathcal{C}} \right] \left[\mathbf{u} \right]_{\mathcal{B}} = \left[f(\mathbf{u}) \right]_{\mathcal{C}}.$$

Proof. **Existence.** Fix $\mathbf{u} \in U$. We must show that

$$\left[\left[f(\mathbf{b}_1) \right]_{\mathcal{C}} \quad \cdots \quad \left[f(\mathbf{b}_m) \right]_{\mathcal{C}} \right] \left[\mathbf{u} \right]_{\mathcal{B}} = \left[f(\mathbf{u}) \right]_{\mathcal{C}}.$$

Set

$$\left[\mathbf{u} \right]_{\mathcal{B}} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix},$$

so that

$$\mathbf{u} = \beta_1 \mathbf{b}_1 + \cdots + \beta_m \mathbf{b}_m.$$

Proof. **Existence.** Fix $\mathbf{u} \in U$. We must show that

$$\left[\ \left[\ f(\mathbf{b}_1)\ \right]_{\mathcal{C}} \quad \ldots \quad \left[\ f(\mathbf{b}_m)\ \right]_{\mathcal{C}}\ \right] \left[\ \mathbf{u}\ \right]_{\mathcal{B}} \;=\; \left[\ f(\mathbf{u})\ \right]_{\mathcal{C}}.$$

Set

$$\left[\ \mathbf{u}\ \right]_{\mathcal{B}} \;=\; \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix},$$

so that

$$\mathbf{u} \;=\; \beta_1 \mathbf{b}_1 + \cdots + \beta_m \mathbf{b}_m.$$

We then compute (next slide):

Proof (continued).

$$
\begin{aligned}
& \left[\left[f(\mathbf{b}_1) \right]_{\mathcal{C}} \quad \dots \quad \left[f(\mathbf{b}_m) \right]_{\mathcal{C}} \right] \left[\mathbf{u} \right]_{\mathcal{B}} \\
= \quad & \left[\left[f(\mathbf{b}_1) \right]_{\mathcal{C}} \quad \dots \quad \left[f(\mathbf{b}_m) \right]_{\mathcal{C}} \right] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \\
= \quad & \beta_1 \left[f(\mathbf{b}_1) \right]_{\mathcal{C}} + \dots + \beta_m \left[f(\mathbf{b}_m) \right]_{\mathcal{C}} \\
\overset{(*)}{=} \quad & \left[\beta_1 f(\mathbf{b}_1) + \dots + \beta_m f(\mathbf{b}_m) \right]_{\mathcal{C}} \\
\overset{(**)}{=} \quad & \left[f\big(\beta_1 \mathbf{b}_1 + \dots + \beta_m \mathbf{b}_m\big) \right]_{\mathcal{C}} \\
= \quad & \left[f(\mathbf{u}) \right]_{\mathcal{C}},
\end{aligned}
$$

where (*) follows from the fact that $\left[\cdot \right]_{\mathcal{C}} : V \to \mathbb{F}^n$ is an isomorphism (and in particular, a linear function), and (**) follows from the fact that f is linear.

Proof (continued). **Uniqueness.**

Proof (continued). **Uniqueness.** Fix any matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ in $\mathbb{F}^{n \times m}$ that has the property that for all $\mathbf{u} \in U$, we have that

$$A \begin{bmatrix} \mathbf{u} \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} f(\mathbf{u}) \end{bmatrix}_{\mathcal{C}}.$$

We must show that

$$A = \begin{bmatrix} \begin{bmatrix} f(\mathbf{b}_1) \end{bmatrix}_{\mathcal{C}} & \dots & \begin{bmatrix} f(\mathbf{b}_m) \end{bmatrix}_{\mathcal{C}} \end{bmatrix}.$$

Proof (continued). **Uniqueness.** Fix any matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ in $\mathbb{F}^{n \times m}$ that has the property that for all $\mathbf{u} \in U$, we have that

$$A \left[\ \mathbf{u} \ \right]_{\mathcal{B}} \ = \ \left[\ f(\mathbf{u}) \ \right]_{\mathcal{C}}.$$

We must show that

$$A \ = \ \left[\ \left[\ f(\mathbf{b}_1) \ \right]_{\mathcal{C}} \ \dots \ \left[\ f(\mathbf{b}_m) \ \right]_{\mathcal{C}} \ \right].$$

We prove this by showing that the two matrices have the same corresponding columns, that is, that $\mathbf{a}_i = \left[\ f(\mathbf{b}_i) \ \right]_{\mathcal{C}}$ for all indices $i \in \{1, \dots, m\}$.

Proof (continued). **Uniqueness.** Fix any matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ in $\mathbb{F}^{n \times m}$ that has the property that for all $\mathbf{u} \in U$, we have that

$$A \begin{bmatrix} \mathbf{u} \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} f(\mathbf{u}) \end{bmatrix}_{\mathcal{C}}.$$

We must show that

$$A = \begin{bmatrix} \begin{bmatrix} f(\mathbf{b}_1) \end{bmatrix}_{\mathcal{C}} & \dots & \begin{bmatrix} f(\mathbf{b}_m) \end{bmatrix}_{\mathcal{C}} \end{bmatrix}.$$

We prove this by showing that the two matrices have the same corresponding columns, that is, that $\mathbf{a}_i = \begin{bmatrix} f(\mathbf{b}_i) \end{bmatrix}_{\mathcal{C}}$ for all indices $i \in \{1, \dots, m\}$. Indeed, for all $i \in \{1, \dots, m\}$, we have the following (next slide):

Proof (continued).

$$\begin{array}{rcll} \mathbf{a}_i & = & A\mathbf{e}_i^m & \text{by Proposition 1.4.4} \\ & = & A\left[\ \mathbf{b}_i\ \right]_{\mathcal{B}} & \text{because } \left[\ \mathbf{b}_i\ \right]_{\mathcal{B}} = \mathbf{e}_i^m \text{ (by Proposition 3.2.9)} \\ & = & \left[\ f(\mathbf{b}_i)\ \right]_{\mathcal{C}} & \text{by the choice of } A. \end{array}$$

This proves that $A = \left[\ \left[\ f(\mathbf{b}_1)\ \right]_{\mathcal{C}} \quad \dots \quad \left[\ f(\mathbf{b}_m)\ \right]_{\mathcal{C}}\ \right]$, and we are done. □

Theorem 4.5.1

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function. Then exists a unique matrix in $\mathbb{F}^{n\times m}$, denoted by ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ and called the *matrix of* f *with respect to* $\mathcal{B}$ *and* $\mathcal{C}$, s.t. for all $\mathbf{u} \in U$, we have that

$${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\ \big[\ \mathbf{u}\ \big]_{\mathcal{B}} \ = \ \big[\ f(\mathbf{u})\ \big]_{\mathcal{C}}.$$

Moreover, the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is given by

$${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \ = \ \Big[\ \big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}}\ \ \dots\ \ \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\ \Big].$$

- **Remark:** Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ are bases of U and V, respectively, and that $f : U \to V$ is a linear function, as in Theorem 4.5.1.

- **Remark:** Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ are bases of U and V, respectively, and that $f : U \to V$ is a linear function, as in Theorem 4.5.1.
  - Then the uniqueness part of Theorem 4.5.1 guarantees that if $A \in \mathbb{F}^{n \times m}$ is **any** matrix that satisfies the property that for all $\mathbf{u} \in U$, we have that

    $$A \left[\ \mathbf{u} \ \right]_{\mathcal{B}} \ = \ \left[\ f(\mathbf{u}) \ \right]_{\mathcal{C}},$$

    then we in fact have that $A = {}_{\mathcal{C}}\left[\ f \ \right]_{\mathcal{B}}$.
  - We will use this observation repeatedly.

- Reminder:

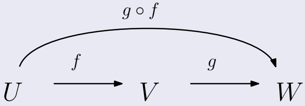
Proposition 4.1.7

Let U, V, and W be vector spaces over a field $\mathbb{F}$. Then all the following hold:

(a) for all linear functions $f, g : U \to V$, the function $f + g$ is linear;

(b) for all linear functions $f : U \to V$ and scalars $\alpha \in \mathbb{F}$, the function $\alpha f : U \to V$ is linear;

(c) for all linear functions $f : U \to V$ and $g : V \to W$, the function $g \circ f$ is linear.



- Reminder:

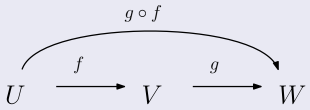
Proposition 4.1.7

Let U, V, and W be vector spaces over a field $\mathbb{F}$. Then all the following hold:

(a) for all linear functions $f, g : U \to V$, the function $f + g$ is linear;

(b) for all linear functions $f : U \to V$ and scalars $\alpha \in \mathbb{F}$, the function $\alpha f : U \to V$ is linear;

(c) for all linear functions $f : U \to V$ and $g : V \to W$, the function $g \circ f$ is linear.



- What about the matrices of sums, scalar multiples, and compositions of linear functions?
  - Next slide!

Theorem 4.5.3

Let U, V, and W be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_p\}$ be a basis of W. Then all the following hold:

(a) for all linear functions $f, g : U \to V$, the function $f + g$ is linear, and moreover,

$$_{\mathcal{C}}[\ f + g\ ]_{\mathcal{B}} \ = \ _{\mathcal{C}}[\ f\ ]_{\mathcal{B}} + {}_{\mathcal{C}}[\ g\ ]_{\mathcal{B}};$$

(b) for all linear functions $f : U \to V$ and scalars $\alpha \in \mathbb{F}$, the function αf is linear, and moreover,

$$_{\mathcal{C}}[\ \alpha f\ ]_{\mathcal{B}} \ = \ \alpha\ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}};$$

(c) for all linear functions $f : U \to V$ and $g : V \to W$, the function $g \circ f$ is linear, and moreover,

$$_{\mathcal{D}}[\ g \circ f\ ]_{\mathcal{B}} \ = \ _{\mathcal{D}}[\ g\ ]_{\mathcal{C}}\ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}.$$

Theorem 4.5.3

Let U, V, and W be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_p\}$ be a basis of W. Then all the following hold:

(a) for all linear functions $f, g : U \to V$, the function $f + g$ is linear, and moreover,

$$ {}_{\mathcal{C}}\big[\ f+g\ \big]_{\mathcal{B}} \ = \ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} + {}_{\mathcal{C}}\big[\ g\ \big]_{\mathcal{B}}; $$

(b) for all linear functions $f : U \to V$ and scalars $\alpha \in \mathbb{F}$, the function αf is linear, and moreover,

$$ {}_{\mathcal{C}}\big[\ \alpha f\ \big]_{\mathcal{B}} \ = \ \alpha\, {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}; $$

(c) for all linear functions $f : U \to V$ and $g : V \to W$, the function $g \circ f$ is linear, and moreover,

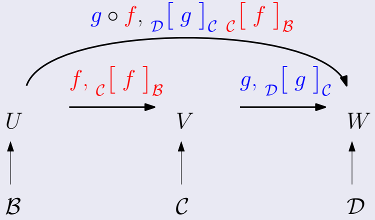
$$ {}_{\mathcal{D}}\big[\ g \circ f\ \big]_{\mathcal{B}} \ = \ {}_{\mathcal{D}}\big[\ g\ \big]_{\mathcal{C}}\ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}. $$

- We prove (c). The proofs of (a) and (b) are left as an exercise.

Theorem 4.5.3

- (c) for all linear functions $f : U \to V$ and $g : V \to W$, the function $g \circ f$ is linear, and moreover,

$$_{\mathcal{D}}[\ g \circ f\ ]_{\mathcal{B}} \ = \ _{\mathcal{D}}[\ g\ ]_{\mathcal{C}}\ _{\mathcal{C}}[\ f\ ]_{\mathcal{B}}.$$

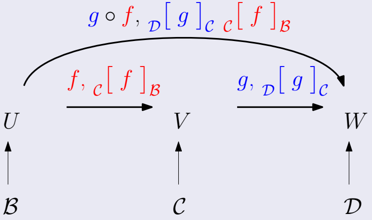


Proof. The fact that $g \circ f$ is linear follows from Proposition 4.1.7(c).

Theorem 4.5.3

(c) for all linear functions $f : U \to V$ and $g : V \to W$, the function $g \circ f$ is linear, and moreover,

$$ {}_{\mathcal{D}}[\ g \circ f\ ]_{\mathcal{B}} \;=\; {}_{\mathcal{D}}[\ g\ ]_{\mathcal{C}}\ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}\,. $$



Proof. The fact that $g \circ f$ is linear follows from Proposition 4.1.7(c). It remains to show that

$$ {}_{\mathcal{D}}[\ g \circ f\ ]_{\mathcal{B}} \;=\; {}_{\mathcal{D}}[\ g\ ]_{\mathcal{C}}\ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}\,. $$

Proof (continued).

Claim. *For all* $\mathbf{u} \in U$, *we have that*

$$\Big({}_{\mathcal{D}}[\ g\ ]_{\mathcal{C}}\ \ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} \Big)\ [\ \mathbf{u}\ ]_{\mathcal{B}} \ = \ [\ (g \circ f)(\mathbf{u})\ ]_{\mathcal{D}}.$$

Proof of the Claim.

Proof (continued).

Claim. *For all* $\mathbf{u} \in U$*, we have that*

$$\left({}_{\mathcal{D}}\left[\ g\ \right]_{\mathcal{C}} \ {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}} \right) \left[\ \mathbf{u}\ \right]_{\mathcal{B}} \;=\; \left[\ (g \circ f)(\mathbf{u})\ \right]_{\mathcal{D}}.$$

Proof of the Claim. For all $\mathbf{u} \in U$, we have the following:

$$\begin{aligned} \left({}_{\mathcal{D}}\left[\ g\ \right]_{\mathcal{C}} \ {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}} \right) \left[\ \mathbf{u}\ \right]_{\mathcal{B}} &= {}_{\mathcal{D}}\left[\ g\ \right]_{\mathcal{C}} \left({}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}} \ \left[\ \mathbf{u}\ \right]_{\mathcal{B}} \right) \\ &= {}_{\mathcal{D}}\left[\ g\ \right]_{\mathcal{C}} \ \left[\ f(\mathbf{u})\ \right]_{\mathcal{C}} \\ &= \left[\ g\big(f(\mathbf{u})\big)\ \right]_{\mathcal{D}} \\ &= \left[\ (g \circ f)(\mathbf{u})\ \right]_{\mathcal{D}}. \end{aligned}$$

This proves the Claim. ⧫

Proof (continued).

Claim. *For all* $\mathbf{u} \in U$, *we have that*

$$\left({}_{\mathcal{D}}[\ g\ ]_{\mathcal{C}}\ \ {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} \right)\ [\ \mathbf{u}\ ]_{\mathcal{B}} \quad = \quad \big[\ (g \circ f)(\mathbf{u})\ \big]_{\mathcal{D}}.$$

Proof (continued).

Claim. *For all* $\mathbf{u} \in U$, *we have that*

$$\left({}_{\mathcal{D}}\big[\ g\ \big]_{\mathcal{C}}\ \ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \right)\ \big[\ \mathbf{u}\ \big]_{\mathcal{B}} \ = \ \big[\ (g \circ f)(\mathbf{u})\ \big]_{\mathcal{D}}.$$

The Claim and the uniqueness part of Theorem 4.5.1 now imply that

$${}_{\mathcal{D}}\big[\ g \circ f\ \big]_{\mathcal{B}} \ = \ {}_{\mathcal{D}}\big[\ g\ \big]_{\mathcal{C}}\ \ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}},$$

which is what we needed to show. □

Theorem 4.5.3

Let U, V, and W be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_p\}$ be a basis of W. Then all the following hold:

(a) for all linear functions $f, g : U \to V$, the function $f + g$ is linear, and moreover,

$$ {}_{\mathcal{C}}\big[\ f + g\ \big]_{\mathcal{B}} \ = \ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} + {}_{\mathcal{C}}\big[\ g\ \big]_{\mathcal{B}}; $$

(b) for all linear functions $f : U \to V$ and scalars $\alpha \in U$, the function αf is linear, and moreover,

$$ {}_{\mathcal{C}}\big[\ \alpha f\ \big]_{\mathcal{B}} \ = \ \alpha\ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}; $$

(c) for all linear functions $f : U \to V$ and $g : V \to W$, the function $g \circ f$ is linear, and moreover,

$$ {}_{\mathcal{D}}\big[\ g \circ f\ \big]_{\mathcal{B}} \ = \ {}_{\mathcal{D}}\big[\ g\ \big]_{\mathcal{C}}\ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}. $$

- We have already seen that it is possible to use the standard matrix of a linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field) in order to determine various properties of f.

- We have already seen that it is possible to use the standard matrix of a linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field) in order to determine various properties of f.
- Our goal will be to generalize those results to linear functions between arbitrary non-trivial, finite-dimensional vector spaces and the matrices of those linear functions (see Theorem 4.5.4 in a couple of slides).

- We have already seen that it is possible to use the standard matrix of a linear function $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field) in order to determine various properties of f.
- Our goal will be to generalize those results to linear functions between arbitrary non-trivial, finite-dimensional vector spaces and the matrices of those linear functions (see Theorem 4.5.4 in a couple of slides).
- Let's first review those old results (and some relevant definitions).

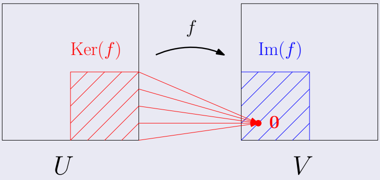
Definition

Given a **linear** function $f : U \to V$, where U and V are vector spaces over a field $\mathbb{F}$, the *kernel* of f is defined to be the set

$$\text{Ker}(f) \quad := \quad \{\mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}\}.$$

The *image* of f is the set

$$\text{Im}(f) \quad := \quad \{f(\mathbf{u}) \mid \mathbf{u} \in U\}.$$



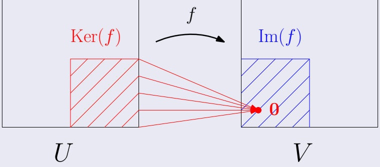
Definition

Given a **linear** function $f : U \to V$, where U and V are vector spaces over a field $\mathbb{F}$, the *kernel* of f is defined to be the set

$$\text{Ker}(f) \quad := \quad \{\mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}\}.$$

The *image* of f is the set

$$\text{Im}(f) \quad := \quad \{f(\mathbf{u}) \mid \mathbf{u} \in U\}.$$



- Reminder: By Theorem 4.2.3, Ker(f) is a subspace of U, and Im(f) is a subspace of V.

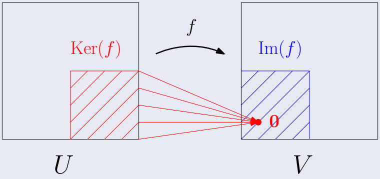
Definition

Given a **linear** function $f : U \to V$, where U and V are vector spaces over a field $\mathbb{F}$, the *kernel* of f is defined to be the set

$$\text{Ker}(f) \ := \ \{\mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}\}.$$

The *image* of f is the set

$$\text{Im}(f) \ := \ \{f(\mathbf{u}) \mid \mathbf{u} \in U\}.$$



- Reminder: By Theorem 4.2.3, $\text{Ker}(f)$ is a subspace of U, and $\text{Im}(f)$ is a subspace of V.
- $\text{rank}(f) := \dim(\text{Im}(f))$

Proposition 4.2.7

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) $\text{rank}(f) = \text{rank}(A)$;

(b) $\dim\big(\text{Ker}(f)\big) = \dim\big(\text{Nul}(A)\big)$.

Proposition 4.2.7

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) $\text{rank}(f) = \text{rank}(A)$;
- (b) $\dim\big(\text{Ker}(f)\big) = \dim\big(\text{Nul}(A)\big)$.

Theorem 1.10.8

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 1.11.9 (abridged)

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a square matrix, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$. Then f is linear and its standard matrix is A. Furthermore, the following are equivalent:

(a) f is an isomorphism;

(b) A is invertible.

Moreover, in this case, f^{-1} is an isomorphism and its standard matrix is A^{-1}.

Theorem 1.11.9 (abridged)

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$ be a square matrix, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$. Then f is linear and its standard matrix is A. Furthermore, the following are equivalent:

(a) f is an isomorphism;

(b) A is invertible.

Moreover, in this case, f^{-1} is an isomorphism and its standard matrix is A^{-1}.

- No matrices, but still relevant to our topic:

Theorem 4.2.4

Let U and V be vector spaces over a field $\mathbb{F}$, and let $f : U \to V$ be a linear function. Then f is one-to-one iff $\text{Ker}(f) = \{\mathbf{0}\}$.

Theorem 1.11.9 (abridged)

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a square matrix, and let $f : \mathbb{F}^n \to \mathbb{F}^n$ be given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$. Then f is linear and its standard matrix is A. Furthermore, the following are equivalent:

(a) f is an isomorphism;

(b) A is invertible.

Moreover, in this case, f^{-1} is an isomorphism and its standard matrix is A^{-1}.

- No matrices, but still relevant to our topic:

Theorem 4.2.4

Let U and V be vector spaces over a field $\mathbb{F}$, and let $f : U \to V$ be a linear function. Then f is one-to-one iff $\text{Ker}(f) = \{\mathbf{0}\}$.

- Now let's generalize these results!

Theorem 4.5.4

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function.[a] Then all the following hold:

- (a) $\text{rank}(f) = \text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)$;
- (b) $\dim\big(\text{Ker}(f)\big) = \dim\Big(\text{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Big)$;
- (c) f is one-to-one iff $\text{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) = \{\mathbf{0}\}$;
- (d) f is one-to-one iff $\text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) = m$ (i.e. the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ has full column rank);
- (e) f is onto iff $\text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) = n$ (i.e. the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ has full row rank);

[a] Note that this means that $\dim(U) = m$, $\dim(V) = n$, and ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \in \mathbb{F}^{n \times m}$.

Theorem 4.5.4 (continued)

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function.[a] Then all the following hold:

(f) f is an isomorphism iff the matrix ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$ is invertible (and in particular, square);

(g) if f is an isomorphism, then ${}_{\mathcal{B}}[\ f^{-1}\ ]_{\mathcal{C}} = \left({}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}\right)^{-1}$.

---

[a] Note that this means that $\dim(U) = m$, $\dim(V) = n$, and ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} \in \mathbb{F}^{n \times m}$.

Theorem 4.5.4 (continued)

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function.[a] Then all the following hold:

(f) f is an isomorphism iff the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is invertible (and in particular, square);

(g) if f is an isomorphism, then ${}_{\mathcal{B}}\big[\ f^{-1}\ \big]_{\mathcal{C}} = \Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)^{-1}$.

[a] Note that this means that $\dim(U) = m$, $\dim(V) = n$, and ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \in \mathbb{F}^{n \times m}$.

- The full proof is in the Lecture Notes.
- Here, we prove parts (a), (b), (c).

Theorem 4.5.4

(a) $\operatorname{rank}(f) = \operatorname{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \Big);$

Proof of (a).

Theorem 4.5.4

(a) $\text{rank}(f) = \text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \Big)$;

Proof of (a). By Theorem 4.5.1, we have that

$$ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \quad = \quad \Big[\ \big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}} \quad \dots \quad \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\ \Big]. $$

We now compute:

$$ \begin{aligned} \text{rank}(f) &= \dim\Big(\text{Span}\big(f(\mathbf{b}_1), \dots, f(\mathbf{b}_m)\big)\Big) \\ &= \dim\Bigg(\text{Span}\Big(\big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}}, \dots, \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\Big)\Bigg) \\ &= \dim\Bigg(\text{Col}\Big(\Big[\ \big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}} \quad \dots \quad \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\ \Big]\Big)\Bigg) \\ &= \dim\Bigg(\text{Col}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Bigg) \quad = \quad \text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big). \end{aligned} $$

□

Theorem 4.5.4

(b) $\dim\big(\mathrm{Ker}(f)\big) = \dim\Big(\mathrm{Nul}\Big({}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}\Big)\Big)$;

Proof of (b).

Theorem 4.5.4

(b) $\dim\big(\mathrm{Ker}(f)\big) = \dim\Big(\mathrm{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Big)$;

Proof of (b). We first observe that

$$\begin{aligned}\mathrm{rank}(f) + \dim\big(\mathrm{Ker}(f)\big) \quad &\overset{(*)}{=} \quad \dim(U) \quad = \quad m \\ &\overset{(**)}{=} \quad \mathrm{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) + \dim\Big(\mathrm{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Big)\end{aligned}$$

where (*) follows from the rank-nullity theorem for linear functions, and (**) follows from the rank-nullity theorem for matrices (since ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is an $n \times m$ matrix).

Theorem 4.5.4

(b) $\dim\big(\text{Ker}(f)\big) = \dim\Big(\text{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Big)$;

Proof of (b). We first observe that

$$\begin{aligned} \text{rank}(f) + \dim\big(\text{Ker}(f)\big) &\overset{(*)}{=} \dim(U) \;=\; m \\ &\overset{(**)}{=} \text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) + \dim\Big(\text{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Big) \end{aligned}$$

where (*) follows from the rank-nullity theorem for linear functions, and (**) follows from the rank-nullity theorem for matrices (since ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is an $n \times m$ matrix).

But by (a), we have that $\text{rank}(f) = \text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)$. Therefore, $\dim\big(\text{Ker}(f)\big) = \dim\Big(\text{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Big)$. $\square$

Theorem 4.5.4

(c) f is one-to-one iff $\text{Nul}\left({}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} \right) = \{\mathbf{0}\}$;

Proof of (c).

Theorem 4.5.4

(c) f is one-to-one iff $\operatorname{Nul}\left({}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}\right) = \{\mathbf{0}\}$;

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

$$
\begin{aligned}
f \text{ is one-to-one} \quad & \overset{(*)}{\iff} \quad \operatorname{Ker}(f) = \{\mathbf{0}\} \\
& \iff \quad \dim\left(\operatorname{Ker}(f)\right) = 0 \\
& \overset{(**)}{\iff} \quad \dim\left(\operatorname{Nul}\left({}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}\right)\right) = 0 \\
& \iff \quad \operatorname{Nul}\left({}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}\right) = \{\mathbf{0}\},
\end{aligned}
$$

where (*) follows from Theorem 4.2.4, and (**) follows from part (b). □

Theorem 4.5.4

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function.[a] Then all the following hold:

(a) $\text{rank}(f) = \text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)$;

(b) $\dim\big(\text{Ker}(f)\big) = \dim\Big(\text{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)\Big)$;

(c) f is one-to-one iff $\text{Nul}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) = \{\mathbf{0}\}$;

(d) f is one-to-one iff $\text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) = m$ (i.e. the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ has full column rank);

(e) f is onto iff $\text{rank}\Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big) = n$ (i.e. the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ has full row rank);

---

[a] Note that this means that $\dim(U) = m$, $\dim(V) = n$, and ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \in \mathbb{F}^{n \times m}$.

Theorem 4.5.4 (continued)

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function.[a] Then all the following hold:

(f) f is an isomorphism iff the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is invertible (and in particular, square);

(g) if f is an isomorphism, then ${}_{\mathcal{B}}\big[\ f^{-1}\ \big]_{\mathcal{C}} = \Big({}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\Big)^{-1}$.

---

[a] Note that this means that $\dim(U) = m$, $\dim(V) = n$, and ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \in \mathbb{F}^{n\times m}$.

- Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ is a basis of U, and that $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ is a basis of V.

- Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ is a basis of U, and that $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ is a basis of V.
- By Theorem 4.5.1, to every linear function $f : U \to V$, we can associate a unique matrix $A \in \mathbb{F}^{n \times m}$ (which we denoted by ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$) s.t. for all $\mathbf{u} \in U$, we have that

$$A \big[\ \mathbf{u}\ \big]_{\mathcal{B}} \quad = \quad \big[\ f(\mathbf{u})\ \big]_{\mathcal{C}}.$$

- Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ is a basis of U, and that $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ is a basis of V.
- By Theorem 4.5.1, to every linear function $f : U \to V$, we can associate a unique matrix $A \in \mathbb{F}^{n \times m}$ (which we denoted by ${}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$) s.t. for all $\mathbf{u} \in U$, we have that

$$A \left[\ \mathbf{u}\ \right]_{\mathcal{B}} \quad = \quad \left[\ f(\mathbf{u})\ \right]_{\mathcal{C}}.$$

- How about the converse? Is it true that for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$?

- Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ is a basis of U, and that $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ is a basis of V.
- By Theorem 4.5.1, to every linear function $f : U \to V$, we can associate a unique matrix $A \in \mathbb{F}^{n \times m}$ (which we denoted by ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$) s.t. for all $\mathbf{u} \in U$, we have that

$$A \left[\ \mathbf{u}\ \right]_{\mathcal{B}} \quad = \quad \left[\ f(\mathbf{u})\ \right]_{\mathcal{C}}.$$

- How about the converse? Is it true that for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$?
- As our next proposition shows, this is indeed true, but the proof is not completely obvious: it relies on several different theorems that we have proven so far.

Proposition 4.5.5

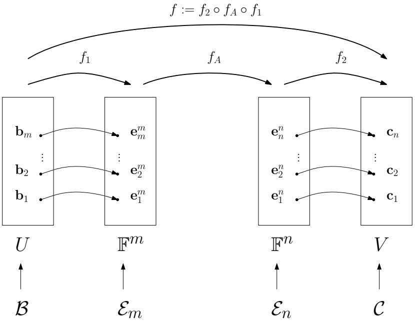
Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$.

Proof (outline).

Proposition 4.5.5

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$.

Proof (outline). **Existence.** The basic idea is in the diagram below. (The full details are in the Lecture Notes.)



Proposition 4.5.5

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$.

Proof (outline, continued). **Uniqueness.**

Proposition 4.5.5

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$.

Proof (outline, continued). **Uniqueness.** Suppose that $f, g : U \to V$ are linear functions s.t. ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} = A$ and ${}_{\mathcal{C}}[\ g\ ]_{\mathcal{B}} = A$. WTS $f = g$.

Proposition 4.5.5

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}$.

Proof (outline, continued). **Uniqueness.** Suppose that $f, g : U \to V$ are linear functions s.t. ${}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}} = A$ and ${}_{\mathcal{C}}[\ g\ ]_{\mathcal{B}} = A$. WTS $f = g$. First of all, note that $\forall i \in \{1, \dots, m\}$:

$$[\ f(\mathbf{b}_i)\ ]_{\mathcal{C}} \;=\; \underbrace{{}_{\mathcal{C}}[\ f\ ]_{\mathcal{B}}}_{=A}\ [\ \mathbf{b}_i\ ]_{\mathcal{B}} \;=\; \underbrace{{}_{\mathcal{C}}[\ g\ ]_{\mathcal{B}}}_{=A}\ [\ \mathbf{b}_i\ ]_{\mathcal{B}} \;=\; [\ g(\mathbf{b}_i)\ ]_{\mathcal{C}},$$

Proposition 4.5.5

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n\times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$.

Proof (outline, continued). **Uniqueness.** Suppose that $f, g : U \to V$ are linear functions s.t. ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} = A$ and ${}_{\mathcal{C}}\big[\ g\ \big]_{\mathcal{B}} = A$. WTS $f = g$. First of all, note that $\forall i \in \{1, \dots, m\}$:

$$\big[\ f(\mathbf{b}_i)\ \big]_{\mathcal{C}} \ = \ \underbrace{{}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}}_{=A} \big[\ \mathbf{b}_i\ \big]_{\mathcal{B}} \ = \ \underbrace{{}_{\mathcal{C}}\big[\ g\ \big]_{\mathcal{B}}}_{=A} \big[\ \mathbf{b}_i\ \big]_{\mathcal{B}} \ = \ \big[\ g(\mathbf{b}_i)\ \big]_{\mathcal{C}},$$

and consequently, $f(\mathbf{b}_i) = g(\mathbf{b}_i)$ (because $\big[\ \cdot\ \big]_{\mathcal{C}} : V \to \mathbb{F}^n$ is an isomorphism and therefore one-to-one).

Proposition 4.5.5

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n \times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$.

Proof (outline, continued). **Uniqueness.** Suppose that $f, g : U \to V$ are linear functions s.t. ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} = A$ and ${}_{\mathcal{C}}\big[\ g\ \big]_{\mathcal{B}} = A$. WTS $f = g$. First of all, note that $\forall i \in \{1, \dots, m\}$:

$$\Big[\ f(\mathbf{b}_i)\ \Big]_{\mathcal{C}} \ = \ \underbrace{{}_{\mathcal{C}}\Big[\ f\ \Big]_{\mathcal{B}}}_{=A} \Big[\ \mathbf{b}_i\ \Big]_{\mathcal{B}} \ = \ \underbrace{{}_{\mathcal{C}}\Big[\ g\ \Big]_{\mathcal{B}}}_{=A} \Big[\ \mathbf{b}_i\ \Big]_{\mathcal{B}} \ = \ \Big[\ g(\mathbf{b}_i)\ \Big]_{\mathcal{C}},$$

and consequently, $f(\mathbf{b}_i) = g(\mathbf{b}_i)$ (because $\big[\ \cdot\ \big]_{\mathcal{C}} : V \to \mathbb{F}^n$ is an isomorphism and therefore one-to-one). But now since $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ is a basis of U and $f, g : U \to V$ are linear, the uniqueness part of Theorem 4.3.2 guarantees that $f = g$. $\square$

Theorem 4.5.1

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function. Then exists a unique matrix in $\mathbb{F}^{n\times m}$, denoted by ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ and called the *matrix of f with respect to $\mathcal{B}$ and $\mathcal{C}$*, s.t. for all $\mathbf{u} \in U$, we have that

$${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}\ \big[\ \mathbf{u}\ \big]_{\mathcal{B}} \ = \ \big[\ f(\mathbf{u})\ \big]_{\mathcal{C}}.$$

Moreover, the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is given by

$${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \ = \ \Big[\ \big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}} \ \ \dots \ \ \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\ \Big].$$

Proposition 4.5.5

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, and let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V. Then for every matrix $A \in \mathbb{F}^{n\times m}$, there exists a unique linear function $f : U \to V$ s.t. $A = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$.

- **Remark:** Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, and recall that $\text{Hom}(U, V)$, the set of all linear functions from U to V, is a vector space over the field $\mathbb{F}$ (vector addition and scalar multiplication in this vector space are the usual addition and scalar multiplication of functions).

- **Remark:** Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, and recall that $\text{Hom}(U, V)$, the set of all linear functions from U to V, is a vector space over the field $\mathbb{F}$ (vector addition and scalar multiplication in this vector space are the usual addition and scalar multiplication of functions).
- Set $m := \dim(U)$ and $n := \dim(V)$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of U and V, respectively.

- **Remark:** Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, and recall that $\text{Hom}(U, V)$, the set of all linear functions from U to V, is a vector space over the field $\mathbb{F}$ (vector addition and scalar multiplication in this vector space are the usual addition and scalar multiplication of functions).
- Set $m := \dim(U)$ and $n := \dim(V)$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of U and V, respectively.
- By Theorem 4.5.1 and Proposition 4.5.5, ${}_{\mathcal{C}}\big[\ \cdot\ \big]_{\mathcal{B}} : \text{Hom}(U, V) \to \mathbb{F}^{n \times m}$ is a bijection, and by Theorem 4.5.3(a-b), it is also a linear function.

- **Remark:** Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, and recall that $\text{Hom}(U, V)$, the set of all linear functions from U to V, is a vector space over the field $\mathbb{F}$ (vector addition and scalar multiplication in this vector space are the usual addition and scalar multiplication of functions).
- Set $m := \dim(U)$ and $n := \dim(V)$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of U and V, respectively.
- By Theorem 4.5.1 and Proposition 4.5.5, ${}_{\mathcal{C}}\big[\ \cdot\ \big]_{\mathcal{B}} : \text{Hom}(U, V) \to \mathbb{F}^{n \times m}$ is a bijection, and by Theorem 4.5.3(a-b), it is also a linear function.
- So, ${}_{\mathcal{C}}\big[\ \cdot\ \big]_{\mathcal{B}} : \text{Hom}(U, V) \to \mathbb{F}^{n \times m}$ is in fact an isomorphism.

- **Remark:** Suppose that U and V are non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, and recall that $\text{Hom}(U, V)$, the set of all linear functions from U to V, is a vector space over the field $\mathbb{F}$ (vector addition and scalar multiplication in this vector space are the usual addition and scalar multiplication of functions).
- Set $m := \dim(U)$ and $n := \dim(V)$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of U and V, respectively.
- By Theorem 4.5.1 and Proposition 4.5.5, ${}_{\mathcal{C}}[\ \cdot\ ]_{\mathcal{B}} : \text{Hom}(U, V) \to \mathbb{F}^{n \times m}$ is a bijection, and by Theorem 4.5.3(a-b), it is also a linear function.
- So, ${}_{\mathcal{C}}[\ \cdot\ ]_{\mathcal{B}} : \text{Hom}(U, V) \to \mathbb{F}^{n \times m}$ is in fact an isomorphism.
- By Theorem 4.2.14(c), it follows that $\dim\Big(\text{Hom}(U, V)\Big) = \dim(\mathbb{F}^{n \times m}) = nm$.

2. Change of basis (transition) matrices

2. Change of basis (transition) matrices

Definition

Given a non-trivial, finite-dimensional vector space V over a field $\mathbb{F}$, and bases $\mathcal{B}$ and $\mathcal{C}$ of V, we call the matrix ${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$ the *change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$* or the *transition matrix from $\mathcal{B}$ to $\mathcal{C}$*.

2. Change of basis (transition) matrices

Definition

Given a non-trivial, finite-dimensional vector space V over a field $\mathbb{F}$, and bases $\mathcal{B}$ and $\mathcal{C}$ of V, we call the matrix ${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$ the *change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$* or the *transition matrix from $\mathcal{B}$ to $\mathcal{C}$*.

Proposition 4.5.6

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases of V. Then the change of basis matrix ${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$ satisfies:

$$ {}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}\ [\ \mathbf{v}\ ]_{\mathcal{B}} \quad = \quad [\ \mathbf{v}\ ]_{\mathcal{C}} \qquad \forall \mathbf{v} \in V. $$

Moreover, this matrix is given by the formula

$$ {}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}} \quad = \quad \left[\ [\ \mathbf{b}_1\ ]_{\mathcal{C}} \quad \ldots \quad [\ \mathbf{b}_n\ ]_{\mathcal{C}}\ \right]. $$

Proposition 4.5.6

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of V. Then the change of basis matrix ${}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$ satisfies:

$$ {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}} \big[\ \mathbf{v}\ \big]_{\mathcal{B}} \ = \ \big[\ \mathbf{v}\ \big]_{\mathcal{C}} \qquad \forall \mathbf{v} \in V. $$

Moreover, this matrix is given by the formula

$$ {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}} \ = \ \Big[\ \big[\ \mathbf{b}_1\ \big]_{\mathcal{C}} \ \ \dots \ \ \big[\ \mathbf{b}_n\ \big]_{\mathcal{C}}\ \Big]. $$

Proof.

Proposition 4.5.6

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of V. Then the change of basis matrix ${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$ satisfies:

$$ {}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}\ [\ \mathbf{v}\ ]_{\mathcal{B}} \ = \ [\ \mathbf{v}\ ]_{\mathcal{C}} \qquad \forall \mathbf{v} \in V. $$

Moreover, this matrix is given by the formula

$$ {}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}} \ = \ \Big[\ [\ \mathbf{b}_1\ ]_{\mathcal{C}} \ \ \dots \ \ [\ \mathbf{b}_n\ ]_{\mathcal{C}}\ \Big]. $$

Proof. The first statement follows straight from the definition of a change of basis matrix; indeed, for all vectors $\mathbf{v} \in V$, we have that

$$ {}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}\ [\ \mathbf{v}\ ]_{\mathcal{B}} \ = \ [\ \mathrm{Id}_V(\mathbf{v})\ ]_{\mathcal{C}} \ = \ [\ \mathbf{v}\ ]_{\mathcal{C}}. $$

Proposition 4.5.6

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of V. Then the change of basis matrix ${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$ satisfies:

$${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}\ [\ \mathbf{v}\ ]_{\mathcal{B}} = [\ \mathbf{v}\ ]_{\mathcal{C}} \qquad \forall \mathbf{v} \in V.$$

Moreover, this matrix is given by the formula

$${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}} = \Big[\ [\ \mathbf{b}_1\ ]_{\mathcal{C}}\ \dots\ [\ \mathbf{b}_n\ ]_{\mathcal{C}}\ \Big].$$

Proof (continued). For the second statement, we observe that

$$\begin{aligned} {}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}} &\overset{(*)}{=} \Big[\ [\ \mathrm{Id}_V(\mathbf{b}_1)\ ]_{\mathcal{C}}\ \dots\ [\ \mathrm{Id}_V(\mathbf{b}_n)\ ]_{\mathcal{C}}\ \Big] \\ &= \Big[\ [\ \mathbf{b}_1\ ]_{\mathcal{C}}\ \dots\ [\ \mathbf{b}_n\ ]_{\mathcal{C}}\ \Big] \end{aligned}$$

where (*) follows from Theorem 4.5.1. □

Proposition 4.5.7

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of V. Then the change of basis matrices ${}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$ and ${}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}$ are invertible, and moreover, they are each other's inverses.

Proof.

Proposition 4.5.7

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of V. Then the change of basis matrices ${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$ and ${}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}$ are invertible, and moreover, they are each other's inverses.

Proof. Clearly, $\mathrm{Id}_V : V \to V$ is an isomorphism, and so by Theorem 4.5.4(f), matrices ${}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$ and ${}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}$ are both invertible. Moreover,

$$ {}_{\mathcal{C}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}} \overset{(*)}{=} {}_{\mathcal{C}}[\ \mathrm{Id}_V^{-1}\ ]_{\mathcal{B}} \overset{(**)}{=} \Big({}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}} \Big)^{-1}, $$

where (*) follows from the fact that $\mathrm{Id}_V^{-1} = \mathrm{Id}_V$, and (**) follows from Theorem 4.5.4(g). This completes the argument. □

- For the special case of $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), we get a nice formula for change of basis matrices (below).

- For the special case of $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), we get a nice formula for change of basis matrices (below).

Theorem 4.5.9

Let $\mathbb{F}$ be a field, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be two bases of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$ and $C := \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_n \end{bmatrix}$. Then the matrix ${}_{\mathcal{C}}\begin{bmatrix} \mathrm{Id}_{\mathbb{F}^n} \end{bmatrix}_{\mathcal{B}}$ is invertible, and it is given by the formula

$$ {}_{\mathcal{C}}\begin{bmatrix} \mathrm{Id}_{\mathbb{F}^n} \end{bmatrix}_{\mathcal{B}} \quad = \quad C^{-1}B. $$

- For the special case of $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), we get a nice formula for change of basis matrices (below).

Theorem 4.5.9

Let $\mathbb{F}$ be a field, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be two bases of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$ and $C := \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_n \end{bmatrix}$. Then the matrix ${}_{\mathcal{C}}\begin{bmatrix} \mathrm{Id}_{\mathbb{F}^n} \end{bmatrix}_{\mathcal{B}}$ is invertible, and it is given by the formula

$$ {}_{\mathcal{C}}\begin{bmatrix} \mathrm{Id}_{\mathbb{F}^n} \end{bmatrix}_{\mathcal{B}} = C^{-1}B. $$

- To prove Theorem 4.5.9, we need a technical lemma.

Lemma 4.5.8

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be any basis of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$. Then B is invertible, and moreover,

$$ {}_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n} \ \big]_{\mathcal{B}} = B \qquad \text{and} \qquad {}_{\mathcal{B}}\big[\ \mathrm{Id}_{\mathbb{F}^n} \ \big]_{\mathcal{E}_n} = B^{-1}. $$

Proof.

Lemma 4.5.8

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be any basis of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$. Then B is invertible, and moreover,

$$_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{B}} \;=\; B \qquad \text{and} \qquad _{\mathcal{B}}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{E}_n} \;=\; B^{-1}.$$

Proof. Let us first prove that $_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{B}} = B$.

Lemma 4.5.8

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be any basis of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$. Then B is invertible, and moreover,

$$ {}_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{B}} = B \qquad \text{and} \qquad {}_{\mathcal{B}}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{E}_n} = B^{-1}. $$

Proof. Let us first prove that ${}_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{B}} = B$. In view of the uniqueness part of Theorem 4.5.1, it suffices to show that $\forall \mathbf{v} \in \mathbb{F}^n$: $B \big[\ \mathbf{v}\ \big]_{\mathcal{B}} = \big[\ \mathbf{v}\ \big]_{\mathcal{E}_n}$.

Lemma 4.5.8

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be any basis of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$. Then B is invertible, and moreover,

$$ {}_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{B}} = B \qquad \text{and} \qquad {}_{\mathcal{B}}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{E}_n} = B^{-1}. $$

Proof. Let us first prove that ${}_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n}\ \big]_{\mathcal{B}} = B$. In view of the uniqueness part of Theorem 4.5.1, it suffices to show that $\forall \mathbf{v} \in \mathbb{F}^n$: $B\big[\ \mathbf{v}\ \big]_{\mathcal{B}} = \big[\ \mathbf{v}\ \big]_{\mathcal{E}_n}$. So, fix a vector $\mathbf{v} \in \mathbb{F}^n$, and set $\big[\ \mathbf{v}\ \big]_{\mathcal{B}} = \begin{bmatrix} \beta_1 & \dots & \beta_n \end{bmatrix}^T$, so that $\mathbf{v} = \beta_1\mathbf{b}_1 + \dots + \beta_n\mathbf{b}_n$.

Lemma 4.5.8

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be any basis of $\mathbb{F}^n$. Set $B := \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{array} \right]$. Then B is invertible, and moreover,

$${}_{\mathcal{E}_n}\left[\text{ Id}_{\mathbb{F}^n} \right]_{\mathcal{B}} \;=\; B \qquad \text{and} \qquad {}_{\mathcal{B}}\left[\text{ Id}_{\mathbb{F}^n} \right]_{\mathcal{E}_n} \;=\; B^{-1}.$$

Proof. Let us first prove that ${}_{\mathcal{E}_n}\left[\text{ Id}_{\mathbb{F}^n} \right]_{\mathcal{B}} = B$. In view of the uniqueness part of Theorem 4.5.1, it suffices to show that $\forall \mathbf{v} \in \mathbb{F}^n$: $B\left[\mathbf{v} \right]_{\mathcal{B}} = \left[\mathbf{v} \right]_{\mathcal{E}_n}$. So, fix a vector $\mathbf{v} \in \mathbb{F}^n$, and set $\left[\mathbf{v} \right]_{\mathcal{B}} = \left[\begin{array}{ccc} \beta_1 & \dots & \beta_n \end{array} \right]^T$, so that $\mathbf{v} = \beta_1\mathbf{b}_1 + \dots + \beta_n\mathbf{b}_n$. Then

$$B\left[\mathbf{v} \right]_{\mathcal{B}} \;=\; \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{array} \right] \left[\begin{array}{c} \beta_1 \\ \vdots \\ \beta_n \end{array} \right] \;=\; \sum_{i=1}^{n} \beta_i \mathbf{b}_i \;=\; \mathbf{v} \;=\; \left[\mathbf{v} \right]_{\mathcal{E}_n}.$$

This proves that ${}_{\mathcal{E}_n}\left[\text{ Id}_{\mathbb{F}^n} \right]_{\mathcal{B}} = B$.

Lemma 4.5.8

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be any basis of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$. Then B is invertible, and moreover,

$$ {}_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n} \ \big]_{\mathcal{B}} = B \qquad \text{and} \qquad {}_{\mathcal{B}}\big[\ \mathrm{Id}_{\mathbb{F}^n} \ \big]_{\mathcal{E}_n} = B^{-1}. $$

Proof (continued). Reminder: ${}_{\mathcal{E}_n}\big[\ \mathrm{Id}_{\mathbb{F}^n} \ \big]_{\mathcal{B}} = B$.

Lemma 4.5.8

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be any basis of $\mathbb{F}^n$. Set $B := \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{array} \right]$. Then B is invertible, and moreover,

$$ {}_{\mathcal{E}_n}\left[\ \mathrm{Id}_{\mathbb{F}^n} \ \right]_{\mathcal{B}} \ = \ B \qquad \text{and} \qquad {}_{\mathcal{B}}\left[\ \mathrm{Id}_{\mathbb{F}^n} \ \right]_{\mathcal{E}_n} \ = \ B^{-1}. $$

Proof (continued). Reminder: ${}_{\mathcal{E}_n}\left[\ \mathrm{Id}_{\mathbb{F}^n} \ \right]_{\mathcal{B}} = B$.

The fact that B is invertible and that ${}_{\mathcal{B}}\left[\ \mathrm{Id}_{\mathbb{F}^n} \ \right]_{\mathcal{E}_n} = B^{-1}$ now follows from Proposition 4.5.7. □

Theorem 4.5.9

Let $\mathbb{F}$ be a field, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be two bases of $\mathbb{F}^n$. Set $B := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{bmatrix}$ and $C := \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_n \end{bmatrix}$. Then the matrix ${}_{\mathcal{C}}\begin{bmatrix} \mathrm{Id}_{\mathbb{F}^n} \end{bmatrix}_{\mathcal{B}}$ is invertible, and it is given by the formula

$$ {}_{\mathcal{C}}\begin{bmatrix} \mathrm{Id}_{\mathbb{F}^n} \end{bmatrix}_{\mathcal{B}} \quad = \quad C^{-1}B. $$

Proof.

Theorem 4.5.9

Let $\mathbb{F}$ be a field, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be two bases of $\mathbb{F}^n$. Set $B := \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{array}\right]$ and $C := \left[\begin{array}{ccc} \mathbf{c}_1 & \dots & \mathbf{c}_n \end{array}\right]$. Then the matrix ${}_{\mathcal{C}}\left[\text{Id}_{\mathbb{F}^n} \right]_{\mathcal{B}}$ is invertible, and it is given by the formula

$$ {}_{\mathcal{C}}\left[\text{Id}_{\mathbb{F}^n} \right]_{\mathcal{B}} = C^{-1}B. $$

Proof. The fact that ${}_{\mathcal{C}}\left[\text{Id}_{\mathbb{F}^n} \right]_{\mathcal{B}}$ is invertible follows from Proposition 4.5.7. To prove that the formula for this matrix is correct, we observe that

$$ \begin{aligned} {}_{\mathcal{C}}\left[\text{Id}_{\mathbb{F}^n} \right]_{\mathcal{B}} = {}_{\mathcal{C}}\left[\text{Id}_{\mathbb{F}^n} \circ \text{Id}_{\mathbb{F}^n} \right]_{\mathcal{B}} &\overset{(*)}{=} {}_{\mathcal{C}}\left[\text{Id}_{\mathbb{F}^n} \right]_{\mathcal{E}_n} \ {}_{\mathcal{E}_n}\left[\text{Id}_{\mathbb{F}^n} \right]_{\mathcal{B}} \\ &\overset{(**)}{=} C^{-1}B, \end{aligned} $$

where (*) follows from Theorem 4.5.3, and (**) follows from Lemma 4.5.8. □

- The following proposition is simply a special case of Theorem 4.5.3(c), but it is used for computation particularly often.

- The following proposition is simply a special case of Theorem 4.5.3(c), but it is used for computation particularly often.

Proposition 4.5.10

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$, let $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases of U, let $\mathcal{C}_1$ and $\mathcal{C}_2$ be bases of V, and let $f : U \to V$ be a linear function. Then

$$
\begin{aligned}
{}_{\mathcal{C}_2}\left[\, f \,\right]_{\mathcal{B}_2} &= {}_{\mathcal{C}_2}\left[\, \mathrm{Id}_V \circ f \circ \mathrm{Id}_U \,\right]_{\mathcal{B}_2} \\
&= {}_{\mathcal{C}_2}\left[\, \mathrm{Id}_V \,\right]_{\mathcal{C}_1} \; {}_{\mathcal{C}_1}\left[\, f \,\right]_{\mathcal{B}_1} \; {}_{\mathcal{B}_1}\left[\, \mathrm{Id}_U \,\right]_{\mathcal{B}_2}
\end{aligned}
$$

Proof. This follows immediately from Theorem 4.5.3(c). □

- Let us now return to the linear function f from Example 4.5.2: we would like to compute its standard matrix.

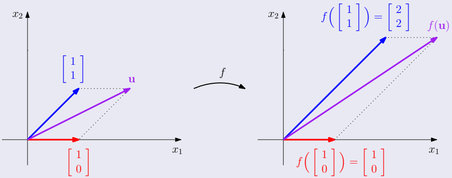
- Let us now return to the linear function f from Example 4.5.2: we would like to compute its standard matrix.

Example 4.5.11

Consider the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ of $\mathbb{R}^2$, and consider the unique linear function $f : \mathbb{R}^2 \to \mathbb{R}^2$ that satisfies the following:

- $f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$;
- $f\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

Compute the standard matrix of the linear function f.



Solution. In Example 4.5.2, we saw that ${}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.
Now, we set $B := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$,[1] and we compute $B^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.
Then the standard matrix of f is

$$\begin{aligned}
{}_{\mathcal{E}_2}\left[\ f\ \right]_{\mathcal{E}_2} &= {}_{\mathcal{E}_2}\left[\ \mathrm{Id}_{\mathbb{R}^2}\ \right]_{\mathcal{B}}\ {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}}\ {}_{\mathcal{B}}\left[\ \mathrm{Id}_{\mathbb{R}^2}\ \right]_{\mathcal{E}_2} && \text{by Prop. 4.5.10} \\
&= B\ {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}}\ B^{-1} && \text{by Lemma 4.5.8} \\
&= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}.
\end{aligned}$$

---

[1] So, the columns of B are the vectors of the basis $\mathcal{B}$, arranged from left to right in the order in which they appear in $\mathcal{B}$.

Solution (continued). Reminder: ${}_{\mathcal{E}_2}[\ f\ ]_{\mathcal{E}_2} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$.

Solution (continued). Reminder: ${}_{\mathcal{E}_2}\big[\ f\ \big]_{\mathcal{E}_2} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$.

Optional: Let us check that our answer is correct. Indeed, we have that

- $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = f\Big(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\Big)$;
- $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = f\Big(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big)$.

So, our answer is correct. □

- Our next proposition essentially states that change of basis matrices are precisely the invertible matrices.

- Our next proposition essentially states that change of basis matrices are precisely the invertible matrices.

Proposition 4.5.12

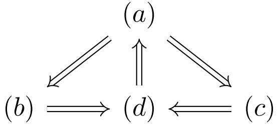
Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a matrix, and let V be any n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) A is invertible;
(b) for all bases $\mathcal{B}$ of V, there exists a basis $\mathcal{C}$ of V s.t. $A = {}_{\mathcal{C}}[\ \text{Id}_V\ ]_{\mathcal{B}}$;
(c) for all bases $\mathcal{C}$ of V, there exists a basis $\mathcal{B}$ of V s.t. $A = {}_{\mathcal{C}}[\ \text{Id}_V\ ]_{\mathcal{B}}$;
(d) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V s.t. $A = {}_{\mathcal{C}}[\ \text{Id}_V\ ]_{\mathcal{B}}$.

Proof. Clearly, it is enough to prove the implications shown in the diagram below.

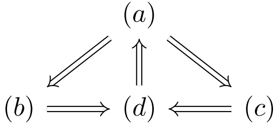
$$\begin{array}{ccccc} & & (a) & & \\ & \swarrow & \Uparrow & \searrow & \\ (b) & \Longrightarrow & (d) & \Longleftarrow & (c) \end{array}$$

Proof. Clearly, it is enough to prove the implications shown in the diagram below.



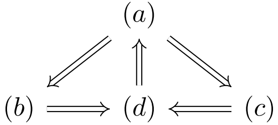
Since V has at least one n-element basis (because $\dim(V) = n$), we see that (b) implies (d), and that (c) implies (d).

Proof. Clearly, it is enough to prove the implications shown in the diagram below.



Since V has at least one n-element basis (because $\dim(V) = n$), we see that (b) implies (d), and that (c) implies (d). Further, by Proposition 4.5.7, (d) implies (a).

Proof. Clearly, it is enough to prove the implications shown in the diagram below.



Since V has at least one n-element basis (because $\dim(V) = n$), we see that (b) implies (d), and that (c) implies (d). Further, by Proposition 4.5.7, (d) implies (a). It remains to show that (a) implies (b) and (c).

Proof. Clearly, it is enough to prove the implications shown in the diagram below.

$$
\begin{array}{ccccc}
 & & (a) & & \\
 & \swarrow & \Uparrow & \searrow & \\
(b) & \Longrightarrow & (d) & \Longleftarrow & (c)
\end{array}
$$

Since V has at least one n-element basis (because $\dim(V) = n$), we see that (b) implies (d), and that (c) implies (d). Further, by Proposition 4.5.7, (d) implies (a). It remains to show that (a) implies (b) and (c). We prove the former; the proof of the latter is similar and is left as an exercise.

Proof (continued). So, assume that (a) is true; we must prove (b).

Proof (continued). So, assume that (a) is true; we must prove (b).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V; we must construct a basis $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ of V s.t. $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$.

Proof (continued). So, assume that (a) is true; we must prove (b).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of V; we must construct a basis $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ of V s.t. $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$.

Using Proposition 4.5.5, we let $f : V \to V$ be the (unique) linear function s.t. $A = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

Proof (continued). So, assume that (a) is true; we must prove (b).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V; we must construct a basis $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ of V s.t. $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$.

Using Proposition 4.5.5, we let $f : V \to V$ be the (unique) linear function s.t. $A = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

Since A is invertible, Theorem 4.5.4(f) guarantees that f is an isomorphism.

Proof (continued). So, assume that (a) is true; we must prove (b).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V; we must construct a basis $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ of V s.t. $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$.

Using Proposition 4.5.5, we let $f : V \to V$ be the (unique) linear function s.t. $A = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

Since A is invertible, Theorem 4.5.4(f) guarantees that f is an isomorphism. Then by Proposition 4.4.1, $f^{-1} : V \to V$ is also an isomorphism.

Proof (continued). So, assume that (a) is true; we must prove (b).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V; we must construct a basis $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ of V s.t. $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$.

Using Proposition 4.5.5, we let $f : V \to V$ be the (unique) linear function s.t. $A = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

Since A is invertible, Theorem 4.5.4(f) guarantees that f is an isomorphism. Then by Proposition 4.4.1, $f^{-1} : V \to V$ is also an isomorphism. For each index $i \in \{1, \dots, n\}$, we set

$$\mathbf{c}_i \quad := \quad f^{-1}(\mathbf{b}_i).$$

Proof (continued). So, assume that (a) is true; we must prove (b).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V; we must construct a basis $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ of V s.t. $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$.

Using Proposition 4.5.5, we let $f : V \to V$ be the (unique) linear function s.t. $A = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

Since A is invertible, Theorem 4.5.4(f) guarantees that f is an isomorphism. Then by Proposition 4.4.1, $f^{-1} : V \to V$ is also an isomorphism. For each index $i \in \{1, \dots, n\}$, we set

$$\mathbf{c}_i \quad := \quad f^{-1}(\mathbf{b}_i).$$

Since $f^{-1} : V \to V$ is an isomorphism and $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis of V, Theorem 4.4.4(c) implies that $\{f^{-1}(\mathbf{b}_1), \dots, f^{-1}(\mathbf{b}_n)\} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\} =: \mathcal{C}$ is also a basis of V.

Proof (continued). Reminder: $A = {}_{\mathcal{C}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{B}}$;
$\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\} = \{f^{-1}(\mathbf{b}_1), \dots, f^{-1}(\mathbf{b}_n)\}$.

Proof (continued). Reminder: $A = {}_{\mathcal{C}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{B}}$;
$\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\} = \{f^{-1}(\mathbf{b}_1), \dots, f^{-1}(\mathbf{b}_n)\}$.

Now, we claim that $A = {}_{\mathcal{C}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{B}}$.

Proof (continued). Reminder: $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$;
$\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\} = \big\{f^{-1}(\mathbf{b}_1), \dots, f^{-1}(\mathbf{b}_n)\big\}$.

Now, we claim that $A = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$. First, we note that

$$\begin{aligned} {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}} &= {}_{\mathcal{C}}\big[\ f^{-1} \circ f\ \big]_{\mathcal{B}} \\ &= {}_{\mathcal{C}}\big[\ f^{-1}\ \big]_{\mathcal{B}}\ \underbrace{{}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}}_{=A} \qquad \text{by Theorem 4.5.3(c)} \\ &= {}_{\mathcal{C}}\big[\ f^{-1}\ \big]_{\mathcal{B}}\ A. \end{aligned}$$

Proof (continued). Reminder: $A = {}_{\mathcal{C}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{B}}$;
$\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\} = \left\{f^{-1}(\mathbf{b}_1), \dots, f^{-1}(\mathbf{b}_n)\right\}$.

Now, we claim that $A = {}_{\mathcal{C}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{B}}$. First, we note that

$$\begin{aligned} {}_{\mathcal{C}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{B}} &= {}_{\mathcal{C}}\left[\ f^{-1} \circ f\ \right]_{\mathcal{B}} \\ &= {}_{\mathcal{C}}\left[\ f^{-1}\ \right]_{\mathcal{B}}\ \underbrace{{}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}}}_{=A} \qquad \text{by Theorem 4.5.3(c)} \\ &= {}_{\mathcal{C}}\left[\ f^{-1}\ \right]_{\mathcal{B}}\ A. \end{aligned}$$

It now suffices to show that ${}_{\mathcal{C}}\left[\ f^{-1}\ \right]_{\mathcal{B}} = I_n$, for it will then immediately follow that $A = {}_{\mathcal{C}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{B}}$, which is what we need.

Proof (continued). Reminder: $A = {}_{\mathcal{C}}\big[\ \text{Id}_V\ \big]_{\mathcal{B}}$;
$\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\} = \{f^{-1}(\mathbf{b}_1), \dots, f^{-1}(\mathbf{b}_n)\}$.

Now, we claim that $A = {}_{\mathcal{C}}\big[\ \text{Id}_V\ \big]_{\mathcal{B}}$. First, we note that

$$\begin{aligned} {}_{\mathcal{C}}\big[\ \text{Id}_V\ \big]_{\mathcal{B}} &= {}_{\mathcal{C}}\big[\ f^{-1} \circ f\ \big]_{\mathcal{B}} \\ &= {}_{\mathcal{C}}\big[\ f^{-1}\ \big]_{\mathcal{B}}\ \underbrace{{}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}}_{=A} \qquad \text{by Theorem 4.5.3(c)} \\ &= {}_{\mathcal{C}}\big[\ f^{-1}\ \big]_{\mathcal{B}}\ A. \end{aligned}$$

It now suffices to show that ${}_{\mathcal{C}}\big[\ f^{-1}\ \big]_{\mathcal{B}} = I_n$, for it will then immediately follow that $A = {}_{\mathcal{C}}\big[\ \text{Id}_V\ \big]_{\mathcal{B}}$, which is what we need. We compute (next slide):

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\} = \{f^{-1}(\mathbf{b}_1), \dots, f^{-1}(\mathbf{b}_n)\}$.

$$\begin{aligned}
{}_{\mathcal{C}}\big[\ f^{-1}\ \big]_{\mathcal{B}} &\overset{(*)}{=} \Big[\ \big[\ f^{-1}(\mathbf{b}_1)\ \big]_{\mathcal{C}} \quad \dots \quad \big[\ f^{-1}(\mathbf{b}_n)\ \big]_{\mathcal{C}}\ \Big] \\
&= \Big[\ \big[\ \mathbf{c}_1\ \big]_{\mathcal{C}} \quad \dots \quad \big[\ \mathbf{c}_n\ \big]_{\mathcal{C}}\ \Big] \\
&\overset{(**)}{=} \big[\ \mathbf{e}_1^n \quad \dots \quad \mathbf{e}_n^n\ \big] \ = \ I_n,
\end{aligned}$$

where (*) follows from Theorem 4.5.1, and (**) follows from Proposition 3.2.9. This proves (b), and we are done. □

3. Similar matrices

3. Similar matrices

Definition

Let $\mathbb{F}$ be a field. Given matrices $A, B \in \mathbb{F}^{n\times n}$, we say that A is *similar* to B if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $B = P^{-1}AP$.

3. Similar matrices

Definition

Let $\mathbb{F}$ be a field. Given matrices $A, B \in \mathbb{F}^{n \times n}$, we say that A is *similar* to B if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

- By Proposition 4.5.13 (below), matrix similarity is an equivalence relation on $\mathbb{F}^{n \times n}$.

3. Similar matrices

Definition

Let $\mathbb{F}$ be a field. Given matrices $A, B \in \mathbb{F}^{n \times n}$, we say that A is *similar* to B if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

- By Proposition 4.5.13 (below), matrix similarity is an equivalence relation on $\mathbb{F}^{n \times n}$.

Proposition 4.5.13

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

(a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;

(b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;

(c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proposition 4.5.13

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

- (a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;
- (b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;
- (c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proof.

Proposition 4.5.13

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

(a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;

(b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;

(c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proof. (a) Fix a matrix $A \in \mathbb{F}^{n \times n}$. Then $A = I_n^{-1} A I_n$, and it follows that A is similar to itself.

Proposition 4.5.13

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

- (a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;
- (b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;
- (c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proof. (a) Fix a matrix $A \in \mathbb{F}^{n \times n}$. Then $A = I_n^{-1} A I_n$, and it follows that A is similar to itself.

(b) Fix a matrices $A, B \in \mathbb{F}^{n \times n}$, and assume that A is similar to B.

Proposition 4.5.13

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

- (a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;
- (b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;
- (c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proof. (a) Fix a matrix $A \in \mathbb{F}^{n \times n}$. Then $A = I_n^{-1} A I_n$, and it follows that A is similar to itself.

(b) Fix a matrices $A, B \in \mathbb{F}^{n \times n}$, and assume that A is similar to B. Then there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1} A P$.

Proposition 4.5.13

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

- (a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;
- (b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;
- (c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proof. (a) Fix a matrix $A \in \mathbb{F}^{n \times n}$. Then $A = I_n^{-1} A I_n$, and it follows that A is similar to itself.

(b) Fix a matrices $A, B \in \mathbb{F}^{n \times n}$, and assume that A is similar to B. Then there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1} A P$. But then $A = PBP^{-1} = (P^{-1})^{-1} B P^{-1}$, and it follows that B is similar to A.

Proposition 4.5.13

(c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proof (continued). (c) Fix matrices $A, B, C \in \mathbb{F}^{n \times n}$, and assume that A is similar to B and that B is similar to C. Then there exist invertible matrices $P, Q \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$ and $C = Q^{-1}BQ$. But now

$$\begin{aligned} C &= Q^{-1}BQ \\ &= Q^{-1}(P^{-1}AP)Q \\ &= (Q^{-1}P^{-1})A(PQ) \\ &= (PQ)^{-1}A(PQ), \end{aligned}$$

and it follows that A is similar to C. □

Proposition 4.5.13

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

(a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;

(b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;

(c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

Proposition 4.5.13

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

(a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;

(b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;

(c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

- **Remark:** By Proposition 4.5.13(b), the similarity relation on $\mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field) is symmetric.

Proposition 4.5.13

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

- (a) $\forall A \in \mathbb{F}^{n\times n}$: A is similar to A;
- (b) $\forall A, B \in \mathbb{F}^{n\times n}$: if A is similar to B, then B is similar to A;
- (c) $\forall A, B, C \in \mathbb{F}^{n\times n}$: if A is similar to B and B is similar to C, then A is similar to C.

- **Remark:** By Proposition 4.5.13(b), the similarity relation on $\mathbb{F}^{n\times n}$ (where $\mathbb{F}$ is a field) is symmetric.
  - Consequently, we may speak of matrices $A, B \in \mathbb{F}^{n\times n}$ as being similar or not being similar **to each other**.

Proposition 4.5.13

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

(a) $\forall A \in \mathbb{F}^{n \times n}$: A is similar to A;

(b) $\forall A, B \in \mathbb{F}^{n \times n}$: if A is similar to B, then B is similar to A;

(c) $\forall A, B, C \in \mathbb{F}^{n \times n}$: if A is similar to B and B is similar to C, then A is similar to C.

- **Remark:** By Proposition 4.5.13(b), the similarity relation on $\mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field) is symmetric.
  - Consequently, we may speak of matrices $A, B \in \mathbb{F}^{n \times n}$ as being similar or not being similar **to each other**.
  - In particular, in what follows, we will often write something like "let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices."
    - This means that A is similar to B and vice versa.

Proposition 4.5.14

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then A is invertible iff B is invertible, and in this case, $B^{-1} = P^{-1}A^{-1}P$ and $A^{-1} = PB^{-1}P^{-1}$.

Proof.

Proposition 4.5.14

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then A is invertible iff B is invertible, and in this case, $B^{-1} = P^{-1}A^{-1}P$ and $A^{-1} = PB^{-1}P^{-1}$.

Proof. Since $B = P^{-1}AP$, we have that $A = PBP^{-1}$. Since P and P^{-1} are invertible, Proposition 1.11.8(e) guarantees that A is invertible iff B is invertible. Suppose now that A and B are invertible. Then

$$B^{-1} = (P^{-1}AP)^{-1} = P^{-1}A^{-1}(P^{-1})^{-1} = P^{-1}A^{-1}P.$$

But now since $B^{-1} = P^{-1}A^{-1}P$, we immediately get that $A^{-1} = PB^{-1}P^{-1}$. This completes the argument. □

Proposition 4.5.14

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then A is invertible iff B is invertible, and in this case, $B^{-1} = P^{-1}A^{-1}P$ and $A^{-1} = PB^{-1}P^{-1}$.

Proposition 4.5.14

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then A is invertible iff B is invertible, and in this case, $B^{-1} = P^{-1}A^{-1}P$ and $A^{-1} = PB^{-1}P^{-1}$.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible,[a] then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

[a] By Proposition 4.5.14, A is invertible iff B is invertible.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^m P$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^m P$ for all integers m.

Proof.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n\times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n\times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof. We first prove that $B^m = P^{-1}A^mP$ for all non-negative integers m.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof. We first prove that $B^m = P^{-1}A^mP$ for all non-negative integers m. We proceed by induction on m.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n\times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n\times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof. We first prove that $B^m = P^{-1}A^mP$ for all non-negative integers m. We proceed by induction on m.

For $m = 0$, we note that $B^0 = I_n$ and $P^{-1}A^0P = P^{-1}I_nP = P^{-1}P = I_n$, and so $B^0 = P^{-1}A^0P$.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof. We first prove that $B^m = P^{-1}A^mP$ for all non-negative integers m. We proceed by induction on m.

For $m = 0$, we note that $B^0 = I_n$ and $P^{-1}A^0P = P^{-1}I_nP = P^{-1}P = I_n$, and so $B^0 = P^{-1}A^0P$.

Now, fix a non-negative integer m, and assume inductively that $B^m = P^{-1}A^mP$. We then have that (next slide):

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof (continued).

$$\begin{aligned} B^{m+1} &= B^m B \overset{\text{ind. hyp.}}{=} (\underbrace{P^{-1}A^mP}_{=B^m})(\underbrace{P^{-1}AP}_{=B}) \\ &= P^{-1}A^m(\underbrace{PP^{-1}}_{=I_n})AP \\ &= P^{-1}A^mAP = P^{-1}A^{m+1}P, \end{aligned}$$

This completes the induction.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof (continued). Reminder: $B^m = P^{-1}A^mP \;\; \forall m \in \mathbb{N}_0$.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof (continued). Reminder: $B^m = P^{-1}A^mP \ \ \forall m \in \mathbb{N}_0$.

Assume now that A and B are invertible.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof (continued). Reminder: $B^m = P^{-1}A^mP \;\; \forall m \in \mathbb{N}_0$.

Assume now that A and B are invertible. By Proposition 4.5.14, we have that $B^{-1} = P^{-1}A^{-1}P$.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n\times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n\times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof (continued). Reminder: $B^m = P^{-1}A^mP \ \ \forall m \in \mathbb{N}_0$.

Assume now that A and B are invertible. By Proposition 4.5.14, we have that $B^{-1} = P^{-1}A^{-1}P$. But now by an argument completely analogous to the above, we get that for all nonegative integers m, we have that $(B^{-1})^m = P^{-1}(A^{-1})^mP$, that is, $B^{-m} = P^{-1}A^{-m}P$.

Proposition 4.5.15

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices, say $B = P^{-1}AP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $B^m = P^{-1}A^mP$, and in particular, A^m and B^m are similar. Moreover, if A and B are invertible, then we in fact have that $B^m = P^{-1}A^mP$ for all integers m.

Proof (continued). Reminder: $B^m = P^{-1}A^mP \;\; \forall m \in \mathbb{N}_0$.

Assume now that A and B are invertible. By Proposition 4.5.14, we have that $B^{-1} = P^{-1}A^{-1}P$. But now by an argument completely analogous to the above, we get that for all nonegative integers m, we have that $(B^{-1})^m = P^{-1}(A^{-1})^mP$, that is, $B^{-m} = P^{-1}A^{-m}P$. Combined with the above, this implies that $B^m = P^{-1}A^mP$ for all integers m. □

- Our next theorem essentially states that two $n \times n$ matrices are similar iff they represent the same linear function from an n-dimensional vector space to itself, but possibly with respect to different bases.

- Our next theorem essentially states that two $n \times n$ matrices are similar iff they represent the same linear function from an n-dimensional vector space to itself, but possibly with respect to different bases.

Theorem 4.5.16

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are similar;

(b) for all bases $\mathcal{B}$ of V and linear functions $f : V \to V$ s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$, there exists a basis $\mathcal{C}$ of V s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$;

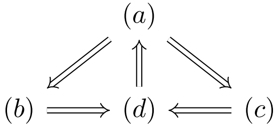
(c) for all bases $\mathcal{C}$ of V and linear functions $f : V \to V$ s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$, there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$;

(d) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$.

Proof. Clearly, it is enough to prove the implications shown in the diagram below.

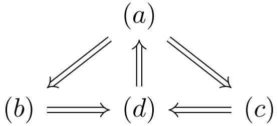
$$
\begin{array}{ccccc}
 & & (a) & & \\
 & \swarrow\!\!\swarrow & \Uparrow & \searrow\!\!\searrow & \\
(b) & \Longrightarrow & (d) & \Longleftarrow & (c)
\end{array}
$$

Proof. Clearly, it is enough to prove the implications shown in the diagram below.



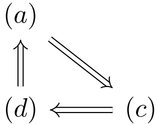
But since matrix similarity in $\mathbb{F}^{n \times n}$ is symmetric (by Proposition 4.5.13(b)), the proofs of the implications "(a) $\Longrightarrow$ (b)" and "(a) $\Longrightarrow$ (c)" are completely analogous, as are the proofs of the implications "(b) $\Longrightarrow$ (d)" and "(c) $\Longrightarrow$ (d)."

Proof. Clearly, it is enough to prove the implications shown in the diagram below.



But since matrix similarity in $\mathbb{F}^{n \times n}$ is symmetric (by Proposition 4.5.13(b)), the proofs of the implications "(a) $\Longrightarrow$ (b)" and "(a) $\Longrightarrow$ (c)" are completely analogous, as are the proofs of the implications "(b) $\Longrightarrow$ (d)" and "(c) $\Longrightarrow$ (d)."

So, it is enough to prove the implications shown in the diagram below.



Proof (continued). First, we assume (a) and prove (c).

Proof (continued). First, we assume (a) and prove (c). Assume that $\mathcal{C}$ is a basis of V and that $f : V \to V$ is a linear function s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$.

Proof (continued). First, we assume (a) and prove (c). Assume that $\mathcal{C}$ is a basis of V and that $f : V \to V$ is a linear function s.t. $C = {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{C}}$. WTS there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}}$.

Proof (continued). First, we assume (a) and prove (c). Assume that $\mathcal{C}$ is a basis of V and that $f : V \to V$ is a linear function s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$. WTS there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$. By (a), matrices B and C are similar, which by definition means that there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}CP$.

Proof (continued). First, we assume (a) and prove (c). Assume that $\mathcal{C}$ is a basis of V and that $f : V \to V$ is a linear function s.t. $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$. WTS there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$. By (a), matrices B and C are similar, which by definition means that there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $B = P^{-1}CP$. Since P is invertible, Proposition 4.5.12 guarantees that there exists a basis $\mathcal{B}$ of V s.t. $P = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$.

Proof (continued). First, we assume (a) and prove (c). Assume that $\mathcal{C}$ is a basis of V and that $f : V \to V$ is a linear function s.t. $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$. WTS there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$. By (a), matrices B and C are similar, which by definition means that there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $B = P^{-1}CP$. Since P is invertible, Proposition 4.5.12 guarantees that there exists a basis $\mathcal{B}$ of V s.t. $P = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$. But now we have that

$$\begin{aligned} B \quad = \quad P^{-1}CP \quad &= \quad \Big({}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}\Big)^{-1}\ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}\ {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}} \\ &\overset{(*)}{=} \quad {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}\ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}\ {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}} \\ &\overset{(**)}{=} \quad {}_{\mathcal{B}}\big[\ \mathrm{Id}_V \circ f \circ \mathrm{Id}_V\ \big]_{\mathcal{B}} \quad = \quad {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}, \end{aligned}$$

where (*) follows from Proposition 4.5.7, and (**) follows from Theorem 4.5.3(c). This proves (c).

Proof (continued). Next, we assume (c) and prove (d).

Proof (continued). Next, we assume (c) and prove (d). Since V is an n-dimensional vector space, it has a basis $\mathcal{C}$ of size n.

Proof (continued). Next, we assume (c) and prove (d). Since V is an n-dimensional vector space, it has a basis $\mathcal{C}$ of size n. Next, by Proposition 4.5.5, there exists a (unique) linear function $f : V \to V$ s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$.

Proof (continued). Next, we assume (c) and prove (d). Since V is an n-dimensional vector space, it has a basis $\mathcal{C}$ of size n. Next, by Proposition 4.5.5, there exists a (unique) linear function $f : V \to V$ s.t. $C = {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{C}}$. But then by (c), there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}}$. This proves (d).

Proof (continued). Next, we assume (c) and prove (d). Since V is an n-dimensional vector space, it has a basis $\mathcal{C}$ of size n. Next, by Proposition 4.5.5, there exists a (unique) linear function $f : V \to V$ s.t. $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$. But then by (c), there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$. This proves (d).

- **Remark:** The implication "(c) $\Longrightarrow$ (d)" may seem trivial, but in fact it is not!
  - To get this implication, we need to make sure that (c) is not just "vacuously true" due to there not existing any $\mathcal{C}$ and f s.t. $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$.
  - The existence of the basis $\mathcal{C}$ follows immediately from dimension considerations, but the existence of a linear function $f : V \to V$ s.t. $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$ only follows from the not entirely trivial Proposition 4.5.5.

Proof (continued). Finally, we assume (d) and prove (a).

Proof (continued). Finally, we assume (d) and prove (a). Using (d), we fix bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$.

Proof (continued). Finally, we assume (d) and prove (a).
Using (d), we fix bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$. Set $P := {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}$.

Proof (continued). Finally, we assume (d) and prove (a). Using (d), we fix bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$. Set $P := {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}$. By Proposition 4.5.7, P is invertible and satisfies $P^{-1} = {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}$. We now compute:

$$\begin{aligned} P^{-1}BP &= {}_{\mathcal{C}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{B}}\ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}} \\ &\overset{(*)}{=} {}_{\mathcal{C}}\big[\ \mathrm{Id}_V \circ f \circ \mathrm{Id}_V\ \big]_{\mathcal{C}} \\ &= {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}} \ = \ C, \end{aligned}$$

where (*) follows from Theorem 4.5.3(c). So, B and C are similar. This proves (a), and we are done. □

Theorem 4.5.16

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are similar;

(b) for all bases $\mathcal{B}$ of V and linear functions $f : V \to V$ s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$, there exists a basis $\mathcal{C}$ of V s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$;

(c) for all bases $\mathcal{C}$ of V and linear functions $f : V \to V$ s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$, there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$;

(d) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$.

Corollary 4.5.17

Let $\mathbb{F}$ be a field, and let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices. Then $\text{rank}(B) = \text{rank}(C)$.

Corollary 4.5.17

Let $\mathbb{F}$ be a field, and let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices. Then $\text{rank}(B) = \text{rank}(C)$.

- This follows immediately from the definition of matrix similarity and from Proposition 3.3.14(c) (below).

Corollary 4.5.17

Let $\mathbb{F}$ be a field, and let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices. Then $\text{rank}(B) = \text{rank}(C)$.

- This follows immediately from the definition of matrix similarity and from Proposition 3.3.14(c) (below).

Proposition 3.3.14

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times m}$. Then all the following hold:

(a) for all invertible matrices $S \in \mathbb{F}^{n \times n}$: $\text{rank}(SA) = \text{rank}(A)$;

(b) for all invertible matrices $S \in \mathbb{F}^{m \times m}$: $\text{rank}(AS) = \text{rank}(A)$;

(c) for all invertible matrices $S_1 \in \mathbb{F}^{n \times n}$ and $S_2 \in \mathbb{F}^{m \times m}$: $\text{rank}(S_1 A S_2) = \text{rank}(A)$.

Corollary 4.5.17

Let $\mathbb{F}$ be a field, and let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices. Then $\text{rank}(B) = \text{rank}(C)$.

- This follows immediately from the definition of matrix similarity and from Proposition 3.3.14(c) (below).

Proposition 3.3.14

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times m}$. Then all the following hold:

(a) for all invertible matrices $S \in \mathbb{F}^{n \times n}$: $\text{rank}(SA) = \text{rank}(A)$;

(b) for all invertible matrices $S \in \mathbb{F}^{m \times m}$: $\text{rank}(AS) = \text{rank}(A)$;

(c) for all invertible matrices $S_1 \in \mathbb{F}^{n \times n}$ and $S_2 \in \mathbb{F}^{m \times m}$: $\text{rank}(S_1 A S_2) = \text{rank}(A)$.

- However, let us give a different proof of Corollary 4.5.17, one relying on Theorem 4.5.16 (in order to illustrate how Theorem 4.5.16 can be used).

Corollary 4.5.17

Let $\mathbb{F}$ be a field, and let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices. Then $\text{rank}(B) = \text{rank}(C)$.

Proof.

Corollary 4.5.17

Let $\mathbb{F}$ be a field, and let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices. Then $\operatorname{rank}(B) = \operatorname{rank}(C)$.

Proof. Since B and C are similar, Theorem 4.5.16 guarantees that there exist bases $\mathcal{B}$ and $\mathcal{C}$ of $\mathbb{F}^n$ and a linear function $f : \mathbb{F}^n \to \mathbb{F}^n$ such that $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$. But then

$$\begin{aligned}
\operatorname{rank}(B) &= \operatorname{rank}\Big({}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}\Big) && \text{because } B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \\
&= \operatorname{rank}(f) && \text{by Theorem 4.5.4(a)} \\
&= \operatorname{rank}\Big({}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}\Big) && \text{by Theorem 4.5.4(a)} \\
&= \operatorname{rank}(C) && \text{because } C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}},
\end{aligned}$$

and we are done. □

4. Checking the existence and uniqueness of linear functions with certain specifications: examples with polynomials and matrices

4. Checking the existence and uniqueness of linear functions with certain specifications: examples with polynomials and matrices

- In what follows, we will see a couple of examples of the following form:

4. Checking the existence and uniqueness of linear functions with certain specifications: examples with polynomials and matrices

- In what follows, we will see a couple of examples of the following form:
    - We are given vector spaces U and V over some field $\mathbb{F}$ (where at least U is non-trivial and finite-dimensional), and we are asked whether there exists a linear function $f : U \to V$ that maps certain specified vectors from U to certain specified vectors from V, and if so, whether this linear function f is unique.

4. Checking the existence and uniqueness of linear functions with certain specifications: examples with polynomials and matrices

- In what follows, we will see a couple of examples of the following form:
  - We are given vector spaces U and V over some field $\mathbb{F}$ (where at least U is non-trivial and finite-dimensional), and we are asked whether there exists a linear function $f : U \to V$ that maps certain specified vectors from U to certain specified vectors from V, and if so, whether this linear function f is unique.
    - In subsection 1.10.4, we already saw such examples for linear functions $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field).
    - Now, we will take a look at examples involving polynomials and matrices (rather than vectors in $\mathbb{F}^n$).

4. Checking the existence and uniqueness of linear functions with certain specifications: examples with polynomials and matrices

- In what follows, we will see a couple of examples of the following form:
  - We are given vector spaces U and V over some field $\mathbb{F}$ (where at least U is non-trivial and finite-dimensional), and we are asked whether there exists a linear function $f : U \to V$ that maps certain specified vectors from U to certain specified vectors from V, and if so, whether this linear function f is unique.
    - In subsection 1.10.4, we already saw such examples for linear functions $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field).
    - Now, we will take a look at examples involving polynomials and matrices (rather than vectors in $\mathbb{F}^n$).
  - We may further be asked to determine various properties of such a linear function f (for example, we may need to determine whether f is an isomorphism).

4. Checking the existence and uniqueness of linear functions with certain specifications: examples with polynomials and matrices

- In what follows, we will see a couple of examples of the following form:
  - We are given vector spaces U and V over some field $\mathbb{F}$ (where at least U is non-trivial and finite-dimensional), and we are asked whether there exists a linear function $f : U \to V$ that maps certain specified vectors from U to certain specified vectors from V, and if so, whether this linear function f is unique.
    - In subsection 1.10.4, we already saw such examples for linear functions $f : \mathbb{F}^m \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field).
    - Now, we will take a look at examples involving polynomials and matrices (rather than vectors in $\mathbb{F}^n$).
  - We may further be asked to determine various properties of such a linear function f (for example, we may need to determine whether f is an isomorphism).
  - In our solutions, we will rely on matrices of linear functions with respect to the most natural bases of the domain and codomain (natural for the vector spaces in question, with no regard to the particular linear function f).

Example 4.5.18

Consider the following matrices with entries in $\mathbb{Z}_2$:

- $M_1 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$;
- $M_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$;
- $M_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$;
- $M_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$;
- $M_5 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$;
- $M_6 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$;
- $M_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$;
- $M_8 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Further, consider the following polynomials with coefficients in $\mathbb{Z}_2$:

- $p_1(x) = x^3 + x^2 + x + 1$;
- $p_2(x) = x^4 + x^2 + x + 1$;
- $p_3(x) = x^5 + x^4 + x^2 + 1$;
- $p_4(x) = x^3$;
- $p_5(x) = x^5 + x^2 + 1$;
- $p_6(x) = x^5 + x^4 + x^2 + x$;
- $p_7(x) = x^2 + x$;
- $p_8(x) = x^5 + x$.

Example 4.5.18 (continued)

1. Prove that there exists a unique linear function $f : \mathbb{Z}_2^{2\times 3} \to \mathbb{P}^5_{\mathbb{Z}_2}$ that satisfies the property that $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$.
2. Find $\text{rank}(f)$ and $\dim(\text{Ker}(f))$.
3. Is f one-to-one? Is it onto? Is it an isomorphism?
4. Find a formula for the linear function f, that is, fill in the blank in the following:

$$f\left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}\right) = \underline{\hspace{6cm}}$$

$$\forall a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}, a_{2,3} \in \mathbb{Z}_2.$$

5. If f is an isomorphism, then find a formula for f^{-1}, that is, fill in the blank in the following:

$$f^{-1}\left(a_5x^5 + \cdots + a_1x + a_0\right) = \underline{\hspace{5cm}}$$

$$\forall a_0, a_1, \ldots, a_5 \in \mathbb{Z}_2.$$

Solution.

Solution. In our solution, we will use the basis

$$\mathcal{M} := \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \right.$$

$$\left. \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

of $\mathbb{Z}_2^{2\times 3}$, and the basis $\mathcal{P} := \{1, x, x^2, x^3, x^4, x^5\}$ of $\mathbb{P}^5_{\mathbb{Z}_2}$.

Example 4.5.18

(a) Prove that there exists a unique linear function $f : \mathbb{Z}_2^{2\times 3} \to \mathbb{P}^5_{\mathbb{Z}_2}$ that satisfies the property that $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$.

Solution (continued). (a)

Example 4.5.18

(a) Prove that there exists a unique linear function $f : \mathbb{Z}_2^{2\times 3} \to \mathbb{P}^5_{\mathbb{Z}_2}$ that satisfies the property that $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$.

Solution (continued). (a) We will solve for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$.

Example 4.5.18

(a) Prove that there exists a unique linear function $f : \mathbb{Z}_2^{2\times 3} \to \mathbb{P}^5_{\mathbb{Z}_2}$ that satisfies the property that $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$.

Solution (continued). (a) We will solve for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$.

We need our linear function f to satisfy $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$, and consequently, our (unknown) matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$ should satisfy

$$ {}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}\ [\ M_i\ ]_{\mathcal{M}} = [\ p_i(x)\ ]_{\mathcal{P}} $$

for all indices $i \in \{1, \ldots, 8\}$.

Example 4.5.18

(a) Prove that there exists a unique linear function $f : \mathbb{Z}_2^{2\times 3} \to \mathbb{P}^5_{\mathbb{Z}_2}$ that satisfies the property that $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$.

Solution (continued). (a) We will solve for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$.

We need our linear function f to satisfy $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$, and consequently, our (unknown) matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$ should satisfy

$$ {}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}} \ [\ M_i\ ]_{\mathcal{M}} \ = \ [\ p_i(x)\ ]_{\mathcal{P}} $$

for all indices $i \in \{1, \ldots, 8\}$. This is equivalent to

$$ {}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}} \underbrace{\left[\ [\ M_1\ ]_{\mathcal{M}} \ \cdots \ [\ M_8\ ]_{\mathcal{M}}\ \right]}_{=:M} = \underbrace{\left[\ [\ p_1(x)\ ]_{\mathcal{P}} \ \cdots \ [\ p_8(x)\ ]_{\mathcal{P}}\ \right]}_{=:P}. $$

Example 4.5.18

(a) Prove that there exists a unique linear function $f : \mathbb{Z}_2^{2\times 3} \to \mathbb{P}^5_{\mathbb{Z}_2}$ that satisfies the property that $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$.

Solution (continued). (a) We will solve for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$.

We need our linear function f to satisfy $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$, and consequently, our (unknown) matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$ should satisfy

$$ {}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}\ [\ M_i\ ]_{\mathcal{M}} \ = \ [\ p_i(x)\ ]_{\mathcal{P}} $$

for all indices $i \in \{1, \ldots, 8\}$. This is equivalent to

$$ {}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}} \underbrace{\left[\ [\ M_1\ ]_{\mathcal{M}} \ \cdots \ [\ M_8\ ]_{\mathcal{M}}\ \right]}_{=:M} \ = \ \underbrace{\left[\ [\ p_1(x)\ ]_{\mathcal{P}} \ \cdots \ [\ p_8(x)\ ]_{\mathcal{P}}\ \right]}_{=:P}. $$

Here, matrices M and P can easily be computed, whereas the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$ is the unknown that we need to solve for.

Solution (continued). Reminder: We need to solve the equation

$$ {}_{\mathcal{P}}\left[\ f\ \right]_{\mathcal{M}} \underbrace{\left[\ \left[\ M_1\ \right]_{\mathcal{M}}\ \ \cdots\ \ \left[\ M_8\ \right]_{\mathcal{M}}\ \right]}_{=:M} = \underbrace{\left[\ \left[\ p_1(x)\ \right]_{\mathcal{P}}\ \ \cdots\ \ \left[\ p_8(x)\ \right]_{\mathcal{P}}\ \right]}_{=:P}. $$

for ${}_{\mathcal{P}}\left[\ f\ \right]_{\mathcal{M}}$.

Solution (continued). Reminder: We need to solve the equation

$$ {}_{\mathcal{P}}\left[\ f\ \right]_{\mathcal{M}} \underbrace{\left[\ \left[\ M_1\ \right]_{\mathcal{M}}\ \ \cdots\ \ \left[\ M_8\ \right]_{\mathcal{M}}\ \right]}_{=:M} \;=\; \underbrace{\left[\ \left[\ p_1(x)\ \right]_{\mathcal{P}}\ \ \cdots\ \ \left[\ p_8(x)\ \right]_{\mathcal{P}}\ \right]}_{=:P}. $$

for ${}_{\mathcal{P}}\left[\ f\ \right]_{\mathcal{M}}$.

We first take the transpose of both sides of the equation above, and we obtain

$$ M^T \left({}_{\mathcal{P}}\left[\ f\ \right]_{\mathcal{M}} \right)^T \;=\; P^T, $$

which we solve for $\left({}_{\mathcal{P}}\left[\ f\ \right]_{\mathcal{M}} \right)^T$.

Solution (continued). We form the matrix

$$\left[\begin{array}{c:c} M^T & P^T \end{array}\right] = \left[\begin{array}{c:c} \left[\begin{array}{c} M_1 \end{array}\right]_{\mathcal{M}}^T & \left[\begin{array}{c} p_1(x) \end{array}\right]_{\mathcal{P}}^T \\ \vdots & \vdots \\ \left[\begin{array}{c} M_8 \end{array}\right]_{\mathcal{M}}^T & \left[\begin{array}{c} p_8(x) \end{array}\right]_{\mathcal{P}}^T \end{array}\right]$$

$$= \left[\begin{array}{cccccc:cccccc} 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right],$$

and we row reduce to obtain (next slide)

Solution.

$$\mathrm{RREF}\Big(\left[\begin{array}{c:c} M^T & P^T \end{array}\right] \Big) \quad = \quad \left[\begin{array}{cccccc:cccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Solution.

$$\mathrm{RREF}\Big(\left[\; M^T \;\vdots\; P^T \;\right] \Big) \;=\; \left[\begin{array}{cccccc:cccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

We now read off the (unique) solution for $\left({}_{\mathcal{P}}[\; f \;]_{\mathcal{M}} \right)^T$:

$$\left({}_{\mathcal{P}}[\; f \;]_{\mathcal{M}} \right)^T \;=\; \left[\begin{array}{cccccc} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

Solution (continued). By taking the transpose, we obtain the (unique) solution for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$:

$$ {}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}} \quad = \quad \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} $$

Solution (continued). By taking the transpose, we obtain the (unique) solution for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$:

$$
{}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}} \quad = \quad \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}
$$

The existence and uniqueness of the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$ guarantees the existence and uniqueness of the linear function $f : \mathbb{Z}_2^{2\times 3} \to \mathbb{P}^5_{\mathbb{Z}_2}$ that satisfies the property that $f(M_i) = p_i(x)$ for all indices $i \in \{1, \ldots, 8\}$.

Solution (continued).

- **Remark:** In the above, the existence and uniqueness of the matrix ${}_{\mathcal{P}}\left[\ f\ \right]_{\mathcal{M}}$ implied the existence and uniqueness of the linear function f with the specifications from the statement of the example.

Solution (continued).

- **Remark:** In the above, the existence and uniqueness of the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$ implied the existence and uniqueness of the linear function f with the specifications from the statement of the example.
  - If we had obtained more than one solution for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$, this would have implied that a linear function f with the given specifications exists, but is not unique.

Solution (continued).

- **Remark:** In the above, the existence and uniqueness of the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$ implied the existence and uniqueness of the linear function f with the specifications from the statement of the example.
  - If we had obtained more than one solution for the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$, this would have implied that a linear function f with the given specifications exists, but is not unique.
  - On the other hand, if there had been no solutions for ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$, this would have meant that no linear function f with the given specifications exists.

Example 4.5.18

(b) Find rank(f) and $\dim(\mathrm{Ker}(f))$.

Solution (continued). (b)

Example 4.5.18

(b) Find $\text{rank}(f)$ and $\dim(\text{Ker}(f))$.

Solution (continued). (b) By row reducing, we see that

$$\text{RREF}\left({}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}} \right) \ = \ I_6.$$

Consequently,

$$\text{rank}(f) \ \overset{(*)}{=} \ \text{rank}\Big({}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}} \Big) \ = \ 6,$$

where (*) follows from Theorem 4.5.4(a).

Example 4.5.18

- (b) Find $\text{rank}(f)$ and $\dim(\text{Ker}(f))$.

Solution (continued). (b) By row reducing, we see that

$$\text{RREF}\left({}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}\right) \;=\; I_6.$$

Consequently,

$$\text{rank}(f) \;\overset{(*)}{=}\; \text{rank}\Big({}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}\Big) \;=\; 6,$$

where (*) follows from Theorem 4.5.4(a).

On the other hand, by the rank-nullity theorem, we have that

$$\text{rank}(f) + \dim(\text{Ker}(f)) \;=\; \dim(\mathbb{Z}_2^{2\times 3}),$$

and it follows that

$$\dim(\text{Ker}(f)) \;=\; \dim(\mathbb{Z}_2^{2\times 3}) - \text{rank}(f) \;=\; 6 - 6 \;=\; 0.$$

Example 4.5.18

(c) Is f one-to-one? Is it onto? Is it an isomorphism?

Solution (continued). (c)

Example 4.5.18

(c) Is f one-to-one? Is it onto? Is it an isomorphism?

Solution (continued). (c) Since $\dim(\mathrm{Ker}(f)) = 0$, Theorem 4.2.4 guarantees that f is one-to-one.

Example 4.5.18

- (c) Is f one-to-one? Is it onto? Is it an isomorphism?

Solution (continued). (c) Since $\dim(\text{Ker}(f)) = 0$, Theorem 4.2.4 guarantees that f is one-to-one.

Since $\text{rank}(f) = 6 = \dim(\mathbb{P}^5_{\mathbb{Z}_2})$, Proposition 4.2.6 guarantees that f is onto.

Example 4.5.18

(c) Is f one-to-one? Is it onto? Is it an isomorphism?

Solution (continued). (c) Since $\dim(\mathrm{Ker}(f)) = 0$, Theorem 4.2.4 guarantees that f is one-to-one.

Since $\mathrm{rank}(f) = 6 = \dim(\mathbb{P}^5_{\mathbb{Z}_2})$, Proposition 4.2.6 guarantees that f is onto.

Since the linear function f is one-to-one and onto, it is an isomorphism.

Example 4.5.18

(c) Is f one-to-one? Is it onto? Is it an isomorphism?

Solution (continued). (c) Since $\dim(\mathrm{Ker}(f)) = 0$, Theorem 4.2.4 guarantees that f is one-to-one.

Since $\mathrm{rank}(f) = 6 = \dim(\mathbb{P}^5_{\mathbb{Z}_2})$, Proposition 4.2.6 guarantees that f is onto.

Since the linear function f is one-to-one and onto, it is an isomorphism.

- Alternatively, since the domain and the codomain of the linear function f have the same finite dimension, and since f is one-to-one, Corollary 4.2.10 guarantees that f is also onto and an isomorphism.

Example 4.5.18

(d) Find a formula for the linear function f, that is, fill in the blank in the following:

$$f\left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}\right) = \underline{\hspace{6cm}}$$

$$\forall a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}, a_{2,3} \in \mathbb{Z}_2.$$

Solution (continued). (d)

Example 4.5.18

(d) Find a formula for the linear function f, that is, fill in the blank in the following:

$$f\left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}\right) = \underline{\hspace{6cm}}$$

$$\forall a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}, a_{2,3} \in \mathbb{Z}_2.$$

Solution (continued). (d) Using the matrix ${}_{\mathcal{P}}[\ f\ ]_{\mathcal{M}}$, we can easily read off the formula for f, as follows.

Solution (continued). For $a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}, a_{2,3} \in \mathbb{Z}_2$:

$$\left[f\left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix} \right) \right]_{\mathcal{P}} = {}_{\mathcal{P}}\big[\ f\ \big]_{\mathcal{M}} \left[\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix} \right]_{\mathcal{M}}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} \\ a_{1,2} \\ a_{1,3} \\ a_{2,1} \\ a_{2,2} \\ a_{2,3} \end{bmatrix}$$

$$= \begin{bmatrix} a_{2,1} \\ a_{2,2} + a_{2,3} \\ a_{2,1} + a_{2,2} \\ a_{1,3} \\ a_{1,1} + a_{2,1} \\ a_{1,1} + a_{1,2} + a_{2,3} \end{bmatrix}$$

$$= \left[\begin{pmatrix} (a_{1,1} + a_{1,2} + a_{2,3})x^5 + \\ +(a_{1,1} + a_{2,1})x^4 + a_{1,3}x^3 + \\ +(a_{2,1} + a_{2,2})x^2 + \\ +(a_{2,2} + a_{2,3})x + a_{2,1} \end{pmatrix} \right]_{\mathcal{P}}$$

Solution (continued). Since $[\ \cdot\ ]_{\mathcal{P}}$ is an isomorphism (and in particular, one-to-one), we deduce that

$$f\Big(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}\Big) = \begin{cases} (a_{1,1} + a_{1,2} + a_{2,3})x^5 + \\ +(a_{1,1} + a_{2,1})x^4 + a_{1,3}x^3 + \\ +(a_{2,1} + a_{2,2})x^2 + \\ +(a_{2,2} + a_{2,3})x + a_{2,1} \end{cases}$$

for all $a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}, a_{2,3} \in \mathbb{Z}_2$. This is the formula that we needed.

Example 4.5.18

(e) If f is an isomorphism, then find a formula for f^{-1}, that is, fill in the blank in the following:

$$f^{-1}\left(a_5x^5 + \cdots + a_1x + a_0\right) = \underline{\hspace{6cm}} \quad \forall a_0, a_1, \ldots, a_5 \in \mathbb{Z}_2.$$

Solution (continued). (e)

Example 4.5.18

(e) If f is an isomorphism, then find a formula for f^{-1}, that is, fill in the blank in the following:

$$f^{-1}\Big(a_5x^5 + \cdots + a_1x + a_0\Big) \quad = \quad \underline{\hspace{6cm}}$$
$$\forall a_0, a_1, \dots, a_5 \in \mathbb{Z}_2.$$

Solution (continued). (e) As we saw in part (c), f is an isomorphism. Let us find a formula for f^{-1}. First, we have that

$$_{\mathcal{M}}\Big[\ f^{-1}\ \Big]_{\mathcal{P}} \overset{(*)}{=} \Big(_{\mathcal{P}}\Big[\ f\ \Big]_{\mathcal{M}}\Big)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix},$$

where (*) follows from Theorem 4.5.4(g).

Example 4.5.18

(e) If f is an isomorphism, then find a formula for f^{-1}, that is, fill in the blank in the following:

$$f^{-1}\Big(a_5x^5 + \cdots + a_1x + a_0\Big) \quad = \quad \underline{\hspace{6cm}}$$

$$\forall a_0, a_1, \dots, a_5 \in \mathbb{Z}_2.$$

Solution (continued). (e) As we saw in part (c), f is an isomorphism. Let us find a formula for f^{-1}. First, we have that

$$_{\mathcal{M}}\big[\ f^{-1}\ \big]_{\mathcal{P}} \overset{(*)}{=} \Big({}_{\mathcal{P}}\big[\ f\ \big]_{\mathcal{M}}\Big)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix},$$

where (*) follows from Theorem 4.5.4(g). We now proceed similarly as in part (d).

Solution (continued). For all $a_0, a_1, a_2, a_3, a_4, a_5 \in \mathbb{Z}_2$, we have the following:

$$\begin{aligned}
& \left[\ f^{-1}\Big(a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0\Big)\ \right]_{\mathcal{M}} \\
= \quad & {}_{\mathcal{M}}\left[\ f^{-1}\ \right]_{\mathcal{P}}\ \left[\ a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0\ \right]_{\mathcal{P}} \\
= \quad & \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \\
= \quad & \begin{bmatrix} a_0 + a_4 \\ a_1 + a_2 + a_4 + a_5 \\ a_3 \\ a_0 \\ a_0 + a_2 \\ a_0 + a_1 + a_2 \end{bmatrix} \\
= \quad & \left[\begin{bmatrix} a_0 + a_4 & a_1 + a_2 + a_4 + a_5 & a_3 \\ a_0 & a_0 + a_2 & a_0 + a_1 + a_2 \end{bmatrix} \right]_{\mathcal{M}}.
\end{aligned}$$

Solution (continued). Since $[\ \cdot\ ]_{\mathcal{M}}$ is an isomorphism (and in particular, one-to-one), it follows that

$$f^{-1}\Big(a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0\Big)$$

$$= \begin{bmatrix} a_0 + a_4 & a_1 + a_2 + a_4 + a_5 & a_3 \\ a_0 & a_0 + a_2 & a_0 + a_1 + a_2 \end{bmatrix}$$

for all $a_0, a_1, a_2, a_3, a_4, a_5 \in \mathbb{Z}_2$. This is the formula for f^{-1} that we needed.

□

Solution (continued). Since $[\ \cdot\ ]_{\mathcal{M}}$ is an isomorphism (and in particular, one-to-one), it follows that

$$f^{-1}\Big(a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0\Big)$$

$$= \begin{bmatrix} a_0 + a_4 & a_1 + a_2 + a_4 + a_5 & a_3 \\ a_0 & a_0 + a_2 & a_0 + a_1 + a_2 \end{bmatrix}$$

for all $a_0, a_1, a_2, a_3, a_4, a_5 \in \mathbb{Z}_2$. This is the formula for f^{-1} that we needed.

- **Optional:** Because it is easy to miscompute, it is a good idea to check our formulas for f and f^{-1}.
  - For f, we do this by plugging in $M_1, \dots, M_8$ into our formula for f (the one that we obtained in part (d)), and checking that we do indeed obtain $p_1(x), \dots, p_8(x)$, resp.
  - Similarly, for f^{-1}, we do this by plugging in $p_1(x), \dots, p_8(x)$ into our formula for f (the one that we obtained in part (e)), and checking that we do indeed obtain $M_1, \dots, M_8$, resp.
  - Details: Lecture Notes. $\square$

Example 4.5.20

Consider the following matrices with entries in $\mathbb{Z}_2$:

- $M_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$;
- $M_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$;
- $M_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$;
- $M_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$;
- $N_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$;
- $N_2 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$;
- $N_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$;
- $N_4 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

Determine if there exists a linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ such that $f(M_i) = N_i$ for all $i \in \{1, 2, 3, 4\}$. If such a linear function f exists, determine if it is unique, and if it is not, determine the number of such linear functions f.

- **Remark:** In this particular case, it is not very hard to see that f does not exist.
  - Indeed, we can see that $M_3 = M_1 + M_2$, and so any linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ satisfying $f(M_1) = N_1$ and $f(M_2) = N_2$ must also satisfy

$$\begin{aligned} f(M_3) &= f(M_1 + M_2) \\ &\stackrel{(*)}{=} f(M_1) + f(M_2) \\ &= N_1 + N_2 \neq N_3, \end{aligned}$$

where (*) follows from the linearity of f.

- **Remark:** In this particular case, it is not very hard to see that f does not exist.
  - Indeed, we can see that $M_3 = M_1 + M_2$, and so any linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ satisfying $f(M_1) = N_1$ and $f(M_2) = N_2$ must also satisfy

$$\begin{aligned} f(M_3) &= f(M_1 + M_2) \\ &\stackrel{(*)}{=} f(M_1) + f(M_2) \\ &= N_1 + N_2 \neq N_3, \end{aligned}$$

    where (*) follows from the linearity of f.
- However, we give a solution that illustrates the general principle, which we can also use in those situations when the non-existence of the function in question is not quite so obvious (and also when the function with the given specifications does in fact exist).

Example 4.5.20

Consider the following matrices with entries in $\mathbb{Z}_2$:

- $M_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$;
- $M_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$;
- $M_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$;
- $M_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$;
- $N_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$;
- $N_2 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$;
- $N_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$;
- $N_4 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

Determine if there exists a linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ such that $f(M_i) = N_i$ for all $i \in \{1, 2, 3, 4\}$. If such a linear function f exists, determine if it is unique, and if it is not, determine the number of such linear functions f.

Solution.

Solution. We proceed as in our solution to Example 4.5.18(a).

Solution. We proceed as in our solution to Example 4.5.18(a). We set

$$A_1 := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A_2 := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad A_3 := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad A_4 := \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

and we further set

$$B_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B_2 := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B_3 := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B_4 := \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad B_5 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad B_6 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In our solution, we will use the basis $\mathcal{A} := \{A_1, A_2, A_3, A_4\}$ of $\mathbb{Z}_2^{2\times 2}$ and the basis $\mathcal{B} := \{B_1, B_2, B_3, B_4, B_5, B_6\}$ of $\mathbb{Z}_2^{2\times 3}$.

Solution (continued). Instead of directly solving for the linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ satisfying $f(M_i) = N_i$ for all $i \in \{1, 2, 3, 4\}$, we will solve for the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}$ in $\mathbb{Z}_2^{6\times 4}$ satisfying

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\ [\ M_i\ ]_{\mathcal{A}} \quad = \quad [\ N_i\ ]_{\mathcal{B}} $$

for all $i \in \{1, 2, 3, 4\}$.

Solution (continued). Instead of directly solving for the linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ satisfying $f(M_i) = N_i$ for all $i \in \{1,2,3,4\}$, we will solve for the matrix ${}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}$ in $\mathbb{Z}_2^{6\times 4}$ satisfying

$$ {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}} \left[\ M_i\ \right]_{\mathcal{A}} \quad = \quad \left[\ N_i\ \right]_{\mathcal{B}} $$

for all $i \in \{1,2,3,4\}$. This is equivalent to

$$ {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}} \underbrace{\left[\ \left[\ M_1\ \right]_{\mathcal{A}} \quad \cdots \quad \left[\ M_4\ \right]_{\mathcal{A}}\ \right]}_{=:M} \quad = \quad \underbrace{\left[\ \left[\ N_1\ \right]_{\mathcal{A}} \quad \cdots \quad \left[\ N_4\ \right]_{\mathcal{A}}\ \right]}_{=:N}. $$

Solution (continued). Instead of directly solving for the linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ satisfying $f(M_i) = N_i$ for all $i \in \{1, 2, 3, 4\}$, we will solve for the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{A}}$ in $\mathbb{Z}_2^{6\times 4}$ satisfying

$$ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{A}} \ \big[\ M_i\ \big]_{\mathcal{A}} \quad = \quad \big[\ N_i\ \big]_{\mathcal{B}} $$

for all $i \in \{1, 2, 3, 4\}$. This is equivalent to

$$ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{A}} \ \underbrace{\Big[\ \big[\ M_1\ \big]_{\mathcal{A}} \ \ \cdots \ \ \big[\ M_4\ \big]_{\mathcal{A}}\ \Big]}_{=:M} \quad = \quad \underbrace{\Big[\ \big[\ N_1\ \big]_{\mathcal{A}} \ \ \cdots \ \ \big[\ N_4\ \big]_{\mathcal{A}}\ \Big]}_{=:N}. $$

Matrices M and N can easily be computed, whereas the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{A}}$ is the unknown that we need to solve for.

Solution (continued). Instead of directly solving for the linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ satisfying $f(M_i) = N_i$ for all $i \in \{1,2,3,4\}$, we will solve for the matrix ${}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}$ in $\mathbb{Z}_2^{6\times 4}$ satisfying

$$ {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}} \left[\ M_i\ \right]_{\mathcal{A}} \quad = \quad \left[\ N_i\ \right]_{\mathcal{B}} $$

for all $i \in \{1,2,3,4\}$. This is equivalent to

$$ {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}} \underbrace{\left[\ \left[\ M_1\ \right]_{\mathcal{A}} \quad \cdots \quad \left[\ M_4\ \right]_{\mathcal{A}}\ \right]}_{=:M} \quad = \quad \underbrace{\left[\ \left[\ N_1\ \right]_{\mathcal{A}} \quad \cdots \quad \left[\ N_4\ \right]_{\mathcal{A}}\ \right]}_{=:N}. $$

Matrices M and N can easily be computed, whereas the matrix ${}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}$ is the unknown that we need to solve for. We first take the transpose of both sides of the equation above, and we obtain

$$ M^T \left({}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}} \right)^T \quad = \quad N^T, $$

which we solve for $\left({}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}} \right)^T$.

Solution (continued). Reminder: We need to solve the equation $M^T \left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}} \right)^T = N^T$ for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}} \right)^T$.

Solution (continued). Reminder: We need to solve the equation $M^T\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T = N^T$ for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T$.

We form the matrix

$$
\left[\begin{array}{c:c} M^T & N^T \end{array}\right] = \left[\begin{array}{c:c} [\ M_1\ ]_{\mathcal{A}}^T & [\ N_1\ ]_{\mathcal{B}}^T \\ \vdots & \vdots \\ [\ M_4\ ]_{\mathcal{A}}^T & [\ N_4\ ]_{\mathcal{B}}^T \end{array}\right]
$$

$$
= \left[\begin{array}{cccc:cccccc} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{array}\right],
$$

Solution (continued). Reminder: We need to solve the equation $M^T \left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}} \right)^T = N^T$ for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}} \right)^T$.

We form the matrix

$$\left[\begin{array}{c:c} M^T & N^T \end{array}\right] = \left[\begin{array}{c:c} [\ M_1\ ]_{\mathcal{A}}^T & [\ N_1\ ]_{\mathcal{B}}^T \\ \vdots & \vdots \\ [\ M_4\ ]_{\mathcal{A}}^T & [\ N_4\ ]_{\mathcal{B}}^T \end{array}\right]$$

$$= \left[\begin{array}{cccc:cccccc} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{array}\right],$$

and we row reduce to obtain

$$\mathrm{RREF}\left(\left[\begin{array}{c:c} M^T & N^T \end{array}\right] \right) = \left[\begin{array}{cccc:cccccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right].$$

Solution (continued). Reminder: We need to solve the equation $M^T\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T = N^T$ for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T$;

$$\mathrm{RREF}\Big(\left[\begin{array}{c;{2pt/2pt}c} M^T & N^T \end{array}\right]\Big) = \left[\begin{array}{cccc;{2pt/2pt}cccccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right].$$

Solution (continued). Reminder: We need to solve the equation $M^T\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T = N^T$ for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T$;

$$\mathrm{RREF}\Big(\left[\begin{array}{c:c} M^T & N^T \end{array}\right]\Big) = \left[\begin{array}{cccc:cccccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right].$$

Because of the fourth row of $\mathrm{RREF}\Big(\left[\begin{array}{c:c} M^T & N^T \end{array}\right]\Big)$, we see that the equation $M^T\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T = N^T$ has no solutions for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T$.

Solution (continued). Reminder: We need to solve the equation $M^T\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T = N^T$ for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T$;

$$\text{RREF}\Big(\left[\begin{array}{c:c} M^T & N^T \end{array}\right]\Big) = \left[\begin{array}{cccc:cccccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right].$$

Because of the fourth row of $\text{RREF}\Big(\left[\begin{array}{c:c} M^T & N^T \end{array}\right]\Big)$, we see that the equation $M^T\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T = N^T$ has no solutions for $\left({}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\right)^T$. Consequently, the equation ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}\ M = N$ has no solutions for ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{A}}$.

Solution (continued). Reminder: We need to solve the equation $M^T\left({}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}\right)^T = N^T$ for $\left({}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}\right)^T$;

$$\mathrm{RREF}\left(\left[\begin{array}{c:c} M^T & N^T \end{array}\right]\right) = \left[\begin{array}{cccc:cccccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right].$$

Because of the fourth row of $\mathrm{RREF}\left(\left[\begin{array}{c:c} M^T & N^T \end{array}\right]\right)$, we see that the equation $M^T\left({}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}\right)^T = N^T$ has no solutions for $\left({}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}\right)^T$. Consequently, the equation ${}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}\ M = N$ has no solutions for ${}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{A}}$.

This implies that there is no linear function $f : \mathbb{Z}_2^{2\times 2} \to \mathbb{Z}_2^{2\times 3}$ satisfying $f(M_i) = N_i$ for all $i \in \{1, 2, 3, 4\}$. $\square$