Linear Algebra 1

Lecture #2

Matrix-vector equations. The rank of a matrix. Matrix operations

Irena Penev

October 16, 2024

- This lecture covers sections 1.4-1.8 of the Lecture Notes (`https://iuuk.mff.cuni.cz/~ipenev/LALectureNotes.pdf`).

This lecture has six parts:

This lecture has six parts:

1. Algebraic operations on vectors and linear span

This lecture has six parts:

1. Algebraic operations on vectors and linear span
2. Matrix-vector multiplication

This lecture has six parts:

1. Algebraic operations on vectors and linear span
2. Matrix-vector multiplication
3. Matrix-vector equations

This lecture has six parts:

1. Algebraic operations on vectors and linear span
2. Matrix-vector multiplication
3. Matrix-vector equations
4. The rank of a matrix

This lecture has six parts:

1. Algebraic operations on vectors and linear span
2. Matrix-vector multiplication
3. Matrix-vector equations
4. The rank of a matrix
5. Matrix operations

This lecture has six parts:

1. Algebraic operations on vectors and linear span
2. Matrix-vector multiplication
3. Matrix-vector equations
4. The rank of a matrix
5. Matrix operations
6. The transpose of a matrix

1. Algebraic operations on vectors and linear span

1. Algebraic operations on vectors and linear span

- Let $\mathbb{F}$ be a field.

1. Algebraic operations on vectors and linear span

- Let $\mathbb{F}$ be a field.
  - For vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ in $\mathbb{F}^n$, we define
    - $\mathbf{x} + \mathbf{y} := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$;
    - $\mathbf{x} - \mathbf{y} := \begin{bmatrix} x_1 - y_1 \\ \vdots \\ x_n - y_n \end{bmatrix}$.

1. Algebraic operations on vectors and linear span

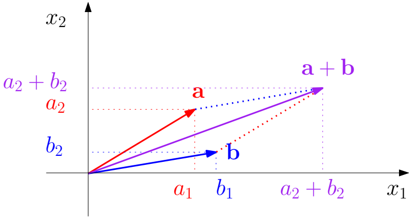
- Let $\mathbb{F}$ be a field.
  - For vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ in $\mathbb{F}^n$, we define
    - $\mathbf{x} + \mathbf{y} := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$;
    - $\mathbf{x} - \mathbf{y} := \begin{bmatrix} x_1 - y_1 \\ \vdots \\ x_n - y_n \end{bmatrix}$.
  - For a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{F}^n$ and a scalar $\alpha \in \mathbb{F}$, we define
    - $\alpha\mathbf{x} := \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$.

Example 1.4.1

Consider the vectors $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}$ in $\mathbb{Z}_3^4$. Then

- $\mathbf{x} + \mathbf{y} = \begin{bmatrix} 0+1 \\ 1+0 \\ 2+2 \\ 2+1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$;
- $2\mathbf{x} = \begin{bmatrix} 2 \cdot 0 \\ 2 \cdot 1 \\ 2 \cdot 2 \\ 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 1 \end{bmatrix}$.

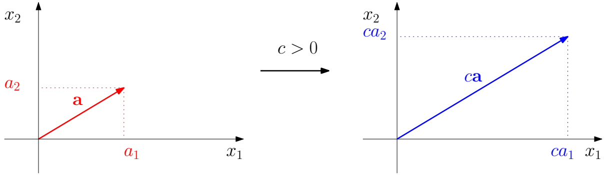
- Vector addition and scalar multiplication in $\mathbb{R}^2$ have a nice geometric interpretation.
- To add two vectors in $\mathbb{R}^2$, say $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, we apply the "parallelogram rule."



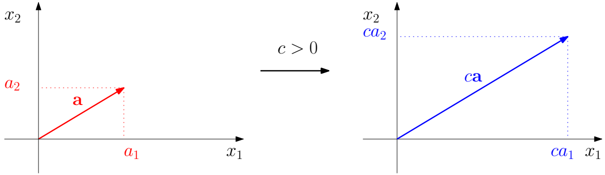
- Scalar multiplication can be interpreted as follows.

- Scalar multiplication can be interpreted as follows.
- Suppose we are given a vector $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and a scalar $c \in \mathbb{R}$.

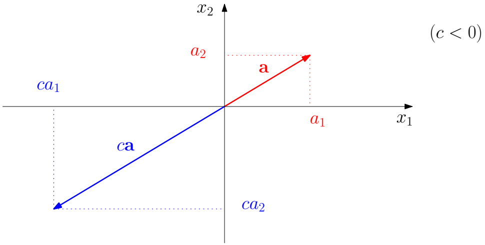
- Scalar multiplication can be interpreted as follows.
- Suppose we are given a vector $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and a scalar $c \in \mathbb{R}$.
- If $c > 0$:



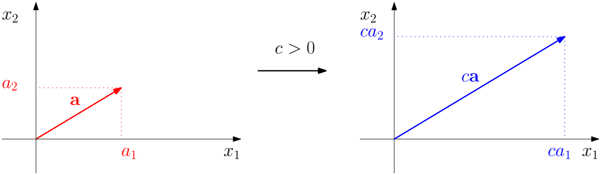
- Scalar multiplication can be interpreted as follows.
- Suppose we are given a vector $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and a scalar $c \in \mathbb{R}$.
- If $c > 0$:



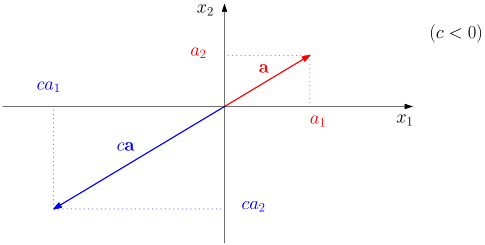
- If $c < 0$:



- Scalar multiplication can be interpreted as follows.
- Suppose we are given a vector $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and a scalar $c \in \mathbb{R}$.
- If $c > 0$:



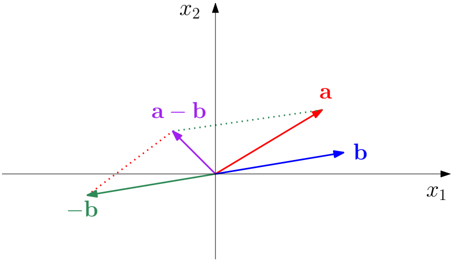
- If $c < 0$:



- If $c = 0$, then $c\mathbf{a} = \mathbf{0}$, which is simply the origin.

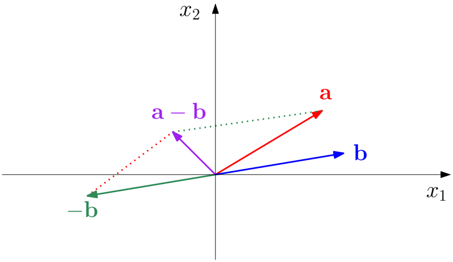
- For vectors $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ in $\mathbb{R}^2$, we note that

$$\mathbf{a} - \mathbf{b} \quad = \quad \mathbf{a} + (-1)\mathbf{b}.$$



- For vectors $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ in $\mathbb{R}^2$, we note that

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-1)\mathbf{b}.$$



- For vectors in $\mathbb{R}^3$, we have a similar geometric interpretation of vector addition and scalar multiplication (and vector subtraction), only in the three-dimensional Euclidean space.

Definition

Suppose $\mathbb{F}$ is some field. A *linear combination* of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ is any sum of the form

$$\sum_{i=1}^{k} \alpha_i \mathbf{v}_i \quad = \quad \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k,$$

where $\alpha_1, \ldots, \alpha_k$ are scalars from the field $\mathbb{F}$.

- For example, in $\mathbb{R}^3$, vectors $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} -3 \\ -9 \\ -3 \end{bmatrix}$ are linear combinations of the vectors $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ because

- For example, in $\mathbb{R}^3$, vectors $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} -3 \\ -9 \\ -3 \end{bmatrix}$ are linear combinations of the vectors $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ because
  - $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$;

- For example, in $\mathbb{R}^3$, vectors $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} -3 \\ -9 \\ -3 \end{bmatrix}$ are linear combinations of the vectors $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ because
  - $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$;
  - $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + (-1) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$;

- For example, in $\mathbb{R}^3$, vectors $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} -3 \\ -9 \\ -3 \end{bmatrix}$ are linear combinations of the vectors $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ because
  - $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$;
  - $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$;
  - $\begin{bmatrix} -3 \\ -9 \\ -3 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = (-3)\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

- For example, in $\mathbb{R}^3$, vectors $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} -3 \\ -9 \\ -3 \end{bmatrix}$ are linear combinations of the vectors $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ because
  - $\begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$;
  - $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + (-1) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$;
  - $\begin{bmatrix} -3 \\ -9 \\ -3 \end{bmatrix} = -3 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = (-3) \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.
- Similarly, $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is a linear combination of the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in $\mathbb{Z}_3^2$ because $\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Definition

Suppose $\mathbb{F}$ is some field. A *linear combination* of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ is any sum of the form

$$\sum_{i=1}^{k} \alpha_i \mathbf{v}_i \quad = \quad \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k,$$

where $\alpha_1, \ldots, \alpha_k$ are scalars from the field $\mathbb{F}$.

Definition

Suppose $\mathbb{F}$ is some field. A *linear combination* of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ is any sum of the form

$$\sum_{i=1}^{k} \alpha_i \mathbf{v}_i = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k,$$

where $\alpha_1, \ldots, \alpha_k$ are scalars from the field $\mathbb{F}$.

- We note that in $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), the zero vector $\mathbf{0}$ is a linear combination of **any** vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ because

$$\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_k.$$

Definition

Suppose $\mathbb{F}$ is some field. A *linear combination* of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ is any sum of the form

$$\sum_{i=1}^{k} \alpha_i \mathbf{v}_i = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k,$$

where $\alpha_1, \ldots, \alpha_k$ are scalars from the field $\mathbb{F}$.

- We note that in $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), the zero vector $\mathbf{0}$ is a linear combination of **any** vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ because

$$\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_k.$$

- Moreover, we define the "empty sum" of vectors in $\mathbb{F}^n$ (or the sum of an "empty list" of vectors in $\mathbb{F}^n$) to be $\mathbf{0}$, where $\mathbf{0}$ is the zero vector in $\mathbb{F}^n$.

Definition

The *linear span* (or simply *span*) of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), denoted by $\mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$ or simply $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, is the set of all linear combinations of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. In other words,

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \;=\; \Big\{ \sum_{i=1}^{k} \alpha_i \mathbf{v}_i \mid \alpha_1, \ldots, \alpha_k \in \mathbb{F} \Big\}.$$

Definition

The *linear span* (or simply *span*) of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), denoted by $\text{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$ or simply $\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, is the set of all linear combinations of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. In other words,

$$\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \Big\{ \sum_{i=1}^{k} \alpha_i \mathbf{v}_i \mid \alpha_1, \ldots, \alpha_k \in \mathbb{F} \Big\}.$$

- So, by definition, a vector $\mathbf{v}$ belongs to $\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ iff it can be written as a linear combination the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Definition

The *linear span* (or simply *span*) of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), denoted by $\mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$ or simply $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, is the set of all linear combinations of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. In other words,

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \quad = \quad \Big\{ \sum_{i=1}^{k} \alpha_i \mathbf{v}_i \mid \alpha_1, \ldots, \alpha_k \in \mathbb{F} \Big\}.$$

- So, by definition, a vector $\mathbf{v}$ belongs to $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ iff it can be written as a linear combination the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.
- As a special case, the empty sum of vectors is equal to the zero vector, and so $\mathrm{Span}(\emptyset) = \{\mathbf{0}\}$.

Definition

The *linear span* (or simply *span*) of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{F}^n$ (where $\mathbb{F}$ is a field), denoted by $\mathrm{Span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$ or simply $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, is the set of all linear combinations of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. In other words,

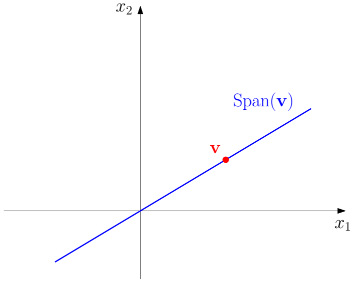
$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \;=\; \Big\{ \sum_{i=1}^{k} \alpha_i \mathbf{v}_i \mid \alpha_1, \ldots, \alpha_k \in \mathbb{F} \Big\}.$$

- So, by definition, a vector $\mathbf{v}$ belongs to $\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ iff it can be written as a linear combination the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.
- As a special case, the empty sum of vectors is equal to the zero vector, and so $\mathrm{Span}(\emptyset) = \{\mathbf{0}\}$.
- Obviously, $\mathrm{Span}(\mathbf{0}) = \{\mathbf{0}\}$.

- Here is a geometric intuition for the special case of $\mathbb{R}^n$.

- Here is a geometric intuition for the special case of $\mathbb{R}^n$.
- As we discussed above, $\text{Span}(\emptyset) = \{\mathbf{0}\}$ and $\text{Span}(\mathbf{0}) = \{\mathbf{0}\}$.

- Here is a geometric intuition for the special case of $\mathbb{R}^n$.
- As we discussed above, $\mathrm{Span}(\emptyset) = \{\mathbf{0}\}$ and $\mathrm{Span}(\mathbf{0}) = \{\mathbf{0}\}$.
- If $\mathbf{v} \neq \mathbf{0}$, then $\mathrm{Span}(\mathbf{v}) = \{\alpha\mathbf{v} \mid \alpha \in \mathbb{R}\}$ is the line through the origin containing $\mathbf{v}$: indeed, $\mathrm{Span}(\mathbf{v})$ is the set of all scalar multiples of $\mathbf{v}$, which is precisely the line through $\mathbf{0}$ and $\mathbf{v}$.



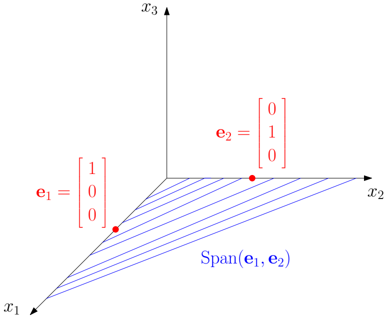
- What if we have two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$?

- What if we have two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$?
- If neither of those vectors is a scalar multiple of the other (and in particular, neither of the two vectors is $\mathbf{0}$), then $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$ is the plane through $\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2$.

- What if we have two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$?
- If neither of those vectors is a scalar multiple of the other (and in particular, neither of the two vectors is $\mathbf{0}$), then $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$ is the plane through $\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2$.
- The case that is particularly easy to visualize is that of the vectors $\mathbf{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ in $\mathbb{R}^3$:

$$
\begin{aligned}
\mathrm{Span}(\mathbf{e}_1, \mathbf{e}_2) &= \Big\{ a_1\mathbf{e}_1 + a_2\mathbf{e}_2 \mid a_1, a_2 \in \mathbb{R} \Big\} \\
&= \left\{ \begin{bmatrix} a_1 \\ a_2 \\ 0 \end{bmatrix} \mid a_1, a_2 \in \mathbb{R} \right\},
\end{aligned}
$$

which is simply the x_1x_2-plane in $\mathbb{R}^3$ (next slide).

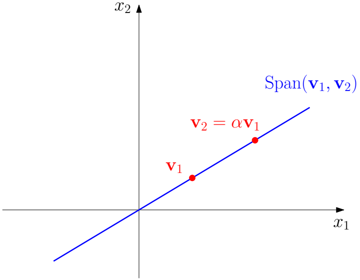


$$\mathrm{Span}(\mathbf{e}_1, \mathbf{e}_2) = \left\{ a_1\mathbf{e}_1 + a_2\mathbf{e}_2 \mid a_1, a_2 \in \mathbb{R} \right\}$$

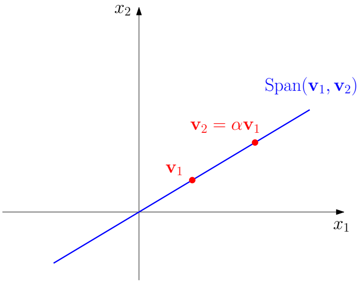
$$= \left\{ \begin{bmatrix} a_1 \\ a_2 \\ 0 \end{bmatrix} \mid a_1, a_2 \in \mathbb{R} \right\}$$

- But what if we have two vectors, one of which is a scalar multiple of the other?

- But what if we have two vectors, one of which is a scalar multiple of the other?
- If $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, with $\mathbf{v}_2 = \alpha \mathbf{v}_1$ for some scalar $\alpha \in \mathbb{R}$ and $\mathbf{v}_1 \neq \mathbf{0}$, then $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$ is the line through the origin, $\mathbf{v}_1$, and $\mathbf{v}_2$.



- But what if we have two vectors, one of which is a scalar multiple of the other?
- If $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, with $\mathbf{v}_2 = \alpha \mathbf{v}_1$ for some scalar $\alpha \in \mathbb{R}$ and $\mathbf{v}_1 \neq \mathbf{0}$, then $\text{Span}(\mathbf{v}_1, \mathbf{v}_2)$ is the line through the origin, $\mathbf{v}_1$, and $\mathbf{v}_2$.



- In general, for vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$, the set $\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is the smallest "flat" (point, line, plane, or higher dimensional generalization) containing the **origin** and all the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

2. Matrix-vector multiplication

2. Matrix-vector multiplication

Definition

Suppose that $\mathbb{F}$ is some field. Given a matrix $A \in \mathbb{F}^{n \times m}$ and a vector $\mathbf{x} \in \mathbb{F}^m$, say

$$A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix},$$

we define the *matrix-vector product* $A\mathbf{x}$ as follows:

$$A\mathbf{x} := \sum_{i=1}^{m} x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \cdots + x_m \mathbf{a}_m.$$

2. Matrix-vector multiplication

Definition

Suppose that $\mathbb{F}$ is some field. Given a matrix $A \in \mathbb{F}^{n \times m}$ and a vector $\mathbf{x} \in \mathbb{F}^m$, say

$$A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix},$$

we define the *matrix-vector product* $A\mathbf{x}$ as follows:

$$A\mathbf{x} := \sum_{i=1}^{m} x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \cdots + x_m \mathbf{a}_m.$$

- Thus, $A\mathbf{x}$ is a linear combination of the columns of A, and the weights/scalars in front of the columns are determined by the entries of the vector $\mathbf{x}$.

- **Reminder:** For a matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \mathbb{F}^{n \times m}$ and vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$ in $\mathbb{F}^m$ (where $\mathbb{F}$ is a field):

$$A\mathbf{x} = \sum_{i=1}^{m} x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \cdots + x_m \mathbf{a}_m.$$

- **Reminder:** For a matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \mathbb{F}^{n \times m}$ and vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$ in $\mathbb{F}^m$ (where $\mathbb{F}$ is a field):

$$A\mathbf{x} = \sum_{i=1}^{m} x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \cdots + x_m \mathbf{a}_m.$$

- Note that, for the matrix-vector product $A\mathbf{x}$ to be defined, two conditions must be satisfied:
  - entries of the matrix A and entries of the vector $\mathbf{x}$ must belong to the same field;
  - the number of **columns** of A must be the same as the number of **entries** of $\mathbf{x}$.

- **Reminder:** For a matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \mathbb{F}^{n \times m}$ and vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$ in $\mathbb{F}^m$ (where $\mathbb{F}$ is a field):

$$A\mathbf{x} = \sum_{i=1}^{m} x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \dots + x_m \mathbf{a}_m.$$

- Note that, for the matrix-vector product $A\mathbf{x}$ to be defined, two conditions must be satisfied:
  - entries of the matrix A and entries of the vector $\mathbf{x}$ must belong to the same field;
  - the number of **columns** of A must be the same as the number of **entries** of $\mathbf{x}$.
- Schematically, we have the following:

$$\underbrace{A}_{\in \mathbb{F}^{n \times m}} \; \underbrace{\mathbf{x}}_{\in \mathbb{F}^{m}} = \underbrace{A\mathbf{x}}_{\in \mathbb{F}^{n}}.$$

Example 1.4.2

Consider the matrix $A \in \mathbb{R}^{3\times 2}$ and vector $\mathbf{x} \in \mathbb{R}^2$, given below:

$$A = \begin{bmatrix} -1 & 2 \\ 2 & 0 \\ 3 & -2 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Then

$$\begin{aligned} A\mathbf{x} &= \begin{bmatrix} -1 & 2 \\ 2 & 0 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= 2 \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 4 \\ 0 \end{bmatrix}. \end{aligned}$$

Example 1.4.3

Consider the matrix $A \in \mathbb{Z}_2^{2\times 3}$ and vector $\mathbf{x} \in \mathbb{Z}_2^3$, given below:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Then

$$\begin{aligned} A\mathbf{x} &= \begin{bmatrix} {\color{red}1} & {\color{blue}1} & {\color{green}0} \\ {\color{red}1} & {\color{blue}0} & {\color{green}1} \end{bmatrix} \begin{bmatrix} {\color{red}1} \\ {\color{blue}1} \\ {\color{green}0} \end{bmatrix} \\ &= {\color{red}1} \begin{bmatrix} {\color{red}1} \\ {\color{red}1} \end{bmatrix} + {\color{blue}1} \begin{bmatrix} {\color{blue}1} \\ {\color{blue}0} \end{bmatrix} + {\color{green}0} \begin{bmatrix} {\color{green}0} \\ {\color{green}1} \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Proposition 1.4.4

Let $\mathbb{F}$ be a field, let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$\mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_m) \quad = \quad \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\}.$$

Proof.

Proposition 1.4.4

Let $\mathbb{F}$ be a field, let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{F}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m) \quad = \quad \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\}.$$

Proof. We compute:

$$\begin{aligned}
\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m) \quad &= \quad \{x_1\mathbf{a}_1 + \dots + x_m\mathbf{a}_m \mid x_1, \dots, x_m \in \mathbb{F}\} \\
&= \quad \left\{ \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \mid x_1, \dots, x_m \in \mathbb{F} \right\} \\
&= \quad \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\}.
\end{aligned}$$

This completes the argument. $\square$

Proposition 1.4.4

Let $\mathbb{F}$ be a field, let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$\mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_m) \quad = \quad \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\}.$$

Proposition 1.4.4

Let $\mathbb{F}$ be a field, let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{F}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m) \;=\; \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\}.$$

- **Remark:** Suppose that $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{F}^n$, where $\mathbb{F}$ is some field, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

Proposition 1.4.4

Let $\mathbb{F}$ be a field, let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{F}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m) \;=\; \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\}.$$

- **Remark:** Suppose that $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{F}^n$, where $\mathbb{F}$ is some field, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.
  - Proposition 1.4.4 states that $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)$, which we defined as the set of all linear combinations of the vectors $\mathbf{a}_1, \dots, \mathbf{a}_m$, is in fact the set of all possible matrix-vector products $A\mathbf{x}$.
    - Here, our matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ is fixed, and the vector $\mathbf{x} \in \mathbb{F}^m$ is allowed to vary.

Proposition 1.4.4

Let $\mathbb{F}$ be a field, let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{F}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m) \quad = \quad \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\}.$$

- **Remark:** Suppose that $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{F}^n$, where $\mathbb{F}$ is some field, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.
  - Proposition 1.4.4 states that $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)$, which we defined as the set of all linear combinations of the vectors $\mathbf{a}_1, \dots, \mathbf{a}_m$, is in fact the set of all possible matrix-vector products $A\mathbf{x}$.
    - Here, our matrix $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ is fixed, and the vector $\mathbf{x} \in \mathbb{F}^m$ is allowed to vary.
  - $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)$, the span of the columns of A, has a special name: it is called the "column space" of the matrix A, and it is denoted by $\text{Col}(A)$.
    - To be studied later in the course.

- Let $\mathbb{F}$ be a field. For each positive integer n and index $i \in \{1, \ldots, n\}$, the vector $\mathbf{e}_i^n$ is the vector in $\mathbb{F}^n$ whose i-th entry is 1, and all of whose other entries are 0's.

$$\mathbf{e}_i^n \quad = \quad \begin{bmatrix} 0 \\ \vdots \\ 0 \\ {\color{red}1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad {\color{red}\longleftarrow \ i\text{-th entry}}$$

- When n is clear from context, we drop the superscript n, and we write $\mathbf{e}_1, \ldots, \mathbf{e}_n$ instead of $\mathbf{e}_1^n, \ldots, \mathbf{e}_n^n$, respectively.
- Vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are called the *standard basis vectors* of $\mathbb{F}^n$, and the set $\mathcal{E}_n := \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is called the *standard basis* of $\mathbb{F}^n$.

$$\mathbf{e}_i^n \quad = \quad \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \longleftarrow i\text{-th entry}$$

- We note that any vector $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{F}^n$ can be expressed as a linear combination of the standard basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in a unique way, namely

$$\mathbf{v} \quad = \quad v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n.$$

Proposition 1.4.5

Let $\mathbb{F}$ be a field, and let $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be a matrix in $\mathbb{F}^{n \times m}$. Then for all indices $i \in \{1, \dots, m\}$, we have that $A\mathbf{e}_i^m = \mathbf{a}_i$.

- **Remark:** Proposition 1.4.4 states that multiplying a matrix by the i-th standard basis vector yields the i-th column of the matrix that we started with.

Proposition 1.4.5

Let $\mathbb{F}$ be a field, and let $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be a matrix in $\mathbb{F}^{n \times m}$. Then for all indices $i \in \{1, \dots, m\}$, we have that $A\mathbf{e}_i^m = \mathbf{a}_i$.

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

$$
\begin{aligned}
A\mathbf{e}_i^m &= \begin{bmatrix} \mathbf{a}_1 & \dots \mathbf{a}_{i-1} & \mathbf{a}_i & \mathbf{a}_{i+1} & \dots & \mathbf{a}_m \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow i\text{-th entry} \\ \\ \\ \\ \end{array} \\
&= 0\mathbf{a}_1 + \dots + 0\mathbf{a}_{i-1} + 1\mathbf{a}_i + 0\mathbf{a}_{i+1} + \dots + 0\mathbf{a}_m \quad = \quad \mathbf{a}_i,
\end{aligned}
$$

which is what we needed to show. $\square$

- For a field $\mathbb{F}$, the *identity matrix* in $\mathbb{F}^{n \times n}$ is the $n \times n$ matrix

$$I_n \quad := \quad \left[\begin{array}{ccc} \mathbf{e}_1^n & \ldots & \mathbf{e}_n^n \end{array} \right].$$

- For a field $\mathbb{F}$, the *identity matrix* in $\mathbb{F}^{n\times n}$ is the $n \times n$ matrix

$$I_n \quad := \quad \left[\begin{array}{ccc} \mathbf{e}_1^n & \dots & \mathbf{e}_n^n \end{array}\right].$$

- In other words, the identity matrix I_n is the $n \times n$ matrix with 1's on the main diagonal and 0's elsewhere (where the 1's and the 0's are from the field $\mathbb{F}$). Schematically, we have that

$$I_n \quad = \quad \left[\begin{array}{cccccc} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{array}\right]_{n\times n}$$

for all positive integers n.

Proposition 1.4.6

Let $\mathbb{F}$ be a field. Then for all vectors $\mathbf{v} \in \mathbb{F}^n$, we have that $I_n\mathbf{v} = \mathbf{v}$.

- **Remark:** Proposition 1.4.5 states that if we multiply the identity matrix by a vector, we obtain that same vector.

For any vector $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{F}^n$, we have that (next slide):

$$
\begin{aligned}
I_n\mathbf{v} &= \begin{bmatrix} \mathbf{e}_1^n & \mathbf{e}_2^n & \ldots & \mathbf{e}_n^n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\
&= v_1\mathbf{e}_1^n + v_2\mathbf{e}_2^n + \cdots + v_n\mathbf{e}_n^n \\
&= v_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + v_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \mathbf{v}. \qquad \square
\end{aligned}
$$

Proposition 1.4.5

Let $\mathbb{F}$ be a field, and let $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$ be a matrix in $\mathbb{F}^{n \times m}$. Then for all indices $i \in \{1, \dots, m\}$, we have that $A\mathbf{e}_i^m = \mathbf{a}_i$.

Proposition 1.4.6

Let $\mathbb{F}$ be a field. Then for all vectors $\mathbf{v} \in \mathbb{F}^n$, we have that $I_n\mathbf{v} = \mathbf{v}$.

Proposition 1.4.5

Let $\mathbb{F}$ be a field, and let $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ be a matrix in $\mathbb{F}^{n \times m}$. Then for all indices $i \in \{1, \dots, m\}$, we have that $A\mathbf{e}_i^m = \mathbf{a}_i$.

Proposition 1.4.6

Let $\mathbb{F}$ be a field. Then for all vectors $\mathbf{v} \in \mathbb{F}^n$, we have that $I_n\mathbf{v} = \mathbf{v}$.

Proposition 1.4.6

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

(a) for all $\mathbf{v} \in \mathbb{F}^m$, we have that $O_{n \times m}\mathbf{v} = \mathbf{0}$;[a]

(b) for all matrices $A \in \mathbb{F}^{n \times m}$, we have that $A\mathbf{0} = \mathbf{0}$.[b]

[a] Here, the zero vector $\mathbf{0}$ belongs to $\mathbb{F}^n$.

[b] In the expression $A\mathbf{0} = \mathbf{0}$, we have that $\mathbf{0} \in \mathbb{F}^m$ and $\mathbf{0} \in \mathbb{F}^n$.

Proof. This readily follows from the definition of matrix-vector multiplication.

3. Matrix-vector equations

- A *matrix-vector equation* is an equation of the form

$$A\mathbf{x} = \mathbf{b},$$

where the matrix A and vector $\mathbf{b}$ are known, and the vector $\mathbf{x}$ is unknown.

3. Matrix-vector equations

- A *matrix-vector equation* is an equation of the form

$$A\mathbf{x} = \mathbf{b},$$

where the matrix A and vector $\mathbf{b}$ are known, and the vector $\mathbf{x}$ is unknown.

  - Here, the entries of A and $\mathbf{b}$ must come from the same field $\mathbb{F}$.

3. Matrix-vector equations

- A *matrix-vector equation* is an equation of the form

$$A\mathbf{x} = \mathbf{b},$$

where the matrix A and vector $\mathbf{b}$ are known, and the vector $\mathbf{x}$ is unknown.
  - Here, the entries of A and $\mathbf{b}$ must come from the same field $\mathbb{F}$.
  - Moreover, the number of **rows** of A must be the same as the number of **entries** of $\mathbf{b}$.

3. Matrix-vector equations

- A *matrix-vector equation* is an equation of the form

$$A\mathbf{x} = \mathbf{b},$$

where the matrix A and vector $\mathbf{b}$ are known, and the vector $\mathbf{x}$ is unknown.

  - Here, the entries of A and $\mathbf{b}$ must come from the same field $\mathbb{F}$.
  - Moreover, the number of **rows** of A must be the same as the number of **entries** of $\mathbf{b}$.
  - Any solution $\mathbf{x}$ will then be a vector in $\mathbb{F}^m$, where m is the number of **columns** of A.

3. Matrix-vector equations

- A *matrix-vector equation* is an equation of the form

$$A\mathbf{x} = \mathbf{b},$$

where the matrix A and vector $\mathbf{b}$ are known, and the vector $\mathbf{x}$ is unknown.
    - Here, the entries of A and $\mathbf{b}$ must come from the same field $\mathbb{F}$.
    - Moreover, the number of **rows** of A must be the same as the number of **entries** of $\mathbf{b}$.
    - Any solution $\mathbf{x}$ will then be a vector in $\mathbb{F}^m$, where m is the number of **columns** of A.
- A matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is equivalent to a system of linear equations whose augmented matrix is $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$.

3. Matrix-vector equations

- A *matrix-vector equation* is an equation of the form

$$A\mathbf{x} = \mathbf{b},$$

where the matrix A and vector $\mathbf{b}$ are known, and the vector $\mathbf{x}$ is unknown.
    - Here, the entries of A and $\mathbf{b}$ must come from the same field $\mathbb{F}$.
    - Moreover, the number of **rows** of A must be the same as the number of **entries** of $\mathbf{b}$.
    - Any solution $\mathbf{x}$ will then be a vector in $\mathbb{F}^m$, where m is the number of **columns** of A.
- A matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is equivalent to a system of linear equations whose augmented matrix is $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$.
    - Details: next slide.

3. Matrix-vector equations

- A *matrix-vector equation* is an equation of the form

$$A\mathbf{x} = \mathbf{b},$$

where the matrix A and vector $\mathbf{b}$ are known, and the vector $\mathbf{x}$ is unknown.
    - Here, the entries of A and $\mathbf{b}$ must come from the same field $\mathbb{F}$.
    - Moreover, the number of **rows** of A must be the same as the number of **entries** of $\mathbf{b}$.
    - Any solution $\mathbf{x}$ will then be a vector in $\mathbb{F}^m$, where m is the number of **columns** of A.
- A matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is equivalent to a system of linear equations whose augmented matrix is $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$.
    - Details: next slide.
    - The matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ will also be referred to as the *augmented matrix* of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$.

$$A\mathbf{x} = \mathbf{b} \iff \underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix}}_{=A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}}_{=\mathbf{x}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_{=\mathbf{b}}$$

$$\iff x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{n,2} \end{bmatrix} + \cdots + x_m \begin{bmatrix} a_{1,m} \\ a_{2,m} \\ \vdots \\ a_{n,m} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\iff \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\iff \left\{ \begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots + a_{1,m}x_m & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots + a_{2,m}x_m & = & b_2 \\ & & & & & \vdots & \\ a_{n,1}x_1 & + & a_{n,2}x_2 & + & \cdots + a_{n,m}x_m & = & b_n \end{array} \right.$$

Example 1.5.1

Solve the matrix-vector equation $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 2 \\ 6 \end{bmatrix},$$

with entries understood to be in $\mathbb{R}$. How many solutions does the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ have?

Solution.

Example 1.5.1

Solve the matrix-vector equation $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 2 \\ 6 \end{bmatrix},$$

with entries understood to be in $\mathbb{R}$. How many solutions does the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ have?

Solution. The augmented matrix of $A\mathbf{x} = \mathbf{b}$ is

$$\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{cc:c} 1 & 2 & 2 \\ 3 & 6 & 6 \end{array}\right].$$

Example 1.5.1

Solve the matrix-vector equation $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 2 \\ 6 \end{bmatrix},$$

with entries understood to be in $\mathbb{R}$. How many solutions does the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ have?

Solution. The augmented matrix of $A\mathbf{x} = \mathbf{b}$ is

$$\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 3 & 6 & 6 \end{array}\right].$$

We now row reduce in order to find $\text{RREF}\left(\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]\right)$, as follows:

$$\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 3 & 6 & 6 \end{array}\right] \overset{R_3 \to R_3 - 3R_1}{\sim} \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 0 & 0 & 0 \end{array}\right].$$

Solution (continued). The last matrix from the computation above is in reduced row echelon form, and we deduce that

$$\mathrm{RREF}\Big(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]\Big) = \left[\begin{array}{cc:c} 1 & 2 & 2 \\ 0 & 0 & 0 \end{array}\right].$$

Solution (continued). The last matrix from the computation above is in reduced row echelon form, and we deduce that

$$\mathrm{RREF}\Big(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] \Big) = \left[\begin{array}{cc:c} 1 & 2 & 2 \\ 0 & 0 & 0 \end{array}\right].$$

The matrix $\mathrm{RREF}(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right])$ is the augmented matrix of the linear system below.

$$\begin{array}{rcrcl} x_1 & + & 2x_2 & = & 2 \\ & & 0 & = & 0 \end{array}$$

Solution (continued). The last matrix from the computation above is in reduced row echelon form, and we deduce that

$$\mathrm{RREF}\Big(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]\Big) = \left[\begin{array}{cc:c} 1 & 2 & 2 \\ 0 & 0 & 0 \end{array}\right].$$

The matrix $\mathrm{RREF}(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right])$ is the augmented matrix of the linear system below.

$$\begin{array}{rcrcr} x_1 & + & 2x_2 & = & 2 \\ & & 0 & = & 0 \end{array}$$

The system is consistent, with one free variable (namely, x_2). We read off the solutions as follows.

$$\begin{array}{rcl} x_1 & = & -2s + 2 \\ x_2 & = & s, \qquad\qquad \text{where } s \in \mathbb{R}. \end{array}$$

Solution (continued). So, the general solution of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} \quad = \quad \begin{bmatrix} -2s + 2 \\ s \end{bmatrix}, \qquad \text{where } s \in \mathbb{R}.$$

Solution (continued). So, the general solution of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = \begin{bmatrix} -2s+2 \\ s \end{bmatrix}, \qquad \text{where } s \in \mathbb{R}.$$

Here is another way to write the general solution of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$:

$$\mathbf{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \qquad \text{where } s \in \mathbb{R}.$$

Solution (continued). So, the general solution of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} \quad = \quad \begin{bmatrix} -2s+2 \\ s \end{bmatrix}, \qquad \text{where } s \in \mathbb{R}.$$

Here is another way to write the general solution of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$:

$$\mathbf{x} \quad = \quad \begin{bmatrix} 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \qquad \text{where } s \in \mathbb{R}.$$

The set of solutions of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\left\{ \begin{bmatrix} -2s+2 \\ s \end{bmatrix} \mid s \in \mathbb{R} \right\} \quad = \quad \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix} \mid s \in \mathbb{R} \right\}.$$

Solution (continued). So, the general solution of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} \;=\; \begin{bmatrix} -2s+2 \\ s \end{bmatrix}, \qquad \text{where } s \in \mathbb{R}.$$

Here is another way to write the general solution of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$:

$$\mathbf{x} \;=\; \begin{bmatrix} 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \qquad \text{where } s \in \mathbb{R}.$$

The set of solutions of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\left\{ \begin{bmatrix} -2s+2 \\ s \end{bmatrix} \mid s \in \mathbb{R} \right\} \;=\; \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix} \mid s \in \mathbb{R} \right\}.$$

Since the parameter s can take infinitely many values (because $\mathbb{R}$ is infinite), the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions. $\square$

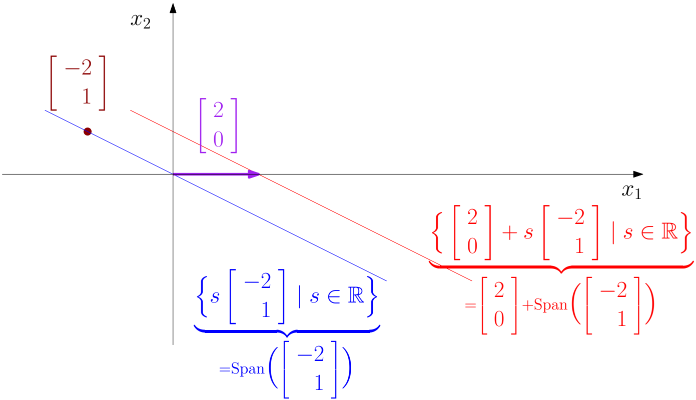
- **Reminder:** The solution set is

$$\left\{ \begin{bmatrix} -2s+2 \\ s \end{bmatrix} \mid s \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix} \mid s \in \mathbb{R} \right\}.$$

- **Reminder:** The solution set is

$$\left\{ \begin{bmatrix} -2s+2 \\ s \end{bmatrix} \;\middle|\; s \in \mathbb{R} \right\} \;=\; \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \end{bmatrix} \;\middle|\; s \in \mathbb{R} \right\}.$$

- This solution set has a geometric interpretation:



Example 1.5.2

Solve the matrix-vector equation $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 2 & 2 & 1 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_3$. How many solutions does the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ have?

Solution.

Example 1.5.2

Solve the matrix-vector equation $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 2 & 2 & 1 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_3$. How many solutions does the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ have?

Solution. The augmented matrix of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 1 & 0 & 1 & 0 & 2 \\ 2 & 2 & 1 & 1 & 0 \end{array}\right].$$

Example 1.5.2

Solve the matrix-vector equation $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 2 & 2 & 1 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_3$. How many solutions does the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ have?

Solution. The augmented matrix of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is

$$\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 1 & 0 & 1 & 0 & 2 \\ 2 & 2 & 1 & 1 & 0 \end{array}\right].$$

We now row reduce in order to find $\text{RREF}(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right])$, as follows (next slide):

Solution (continued).

$$
\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 2 \\ 1 & 0 & 1 & 0 & 2 \\ 2 & 2 & 1 & 1 & 0 \end{array}\right]
$$

$$
\underset{\sim}{\substack{R_2 \to R_2 + 2R_1 \\ R_3 \to R_3 + R_1}} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 & 2 \end{array}\right]
$$

$$
\underset{\sim}{R_3 \to R_3 + 2R_2} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]
$$

$$
\underset{\sim}{R_3 \to 2R_3} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]
$$

$$
\underset{\sim}{R_1 \to R_1 + R_3} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]
$$

$$
\underset{\sim}{R_1 \to R_1 + R_2} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].
$$

Solution (continued). The last matrix from the computation above is in reduced row echelon form, and we deduce that

$$\text{RREF}\Big(\left[\begin{array}{c:c} A & \mathbf{b} \end{array} \right] \Big) \quad = \quad \left[\begin{array}{cccc:c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

Solution (continued). The last matrix from the computation above is in reduced row echelon form, and we deduce that

$$\mathsf{RREF}\Big(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]\Big) \;=\; \left[\begin{array}{cccc:c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

We see from $\mathsf{RREF}(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right])$ that the rightmost column of $\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]$ is a pivot column; consequently, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is inconsistent, i.e. the solution set of the equation $A\mathbf{x} = \mathbf{b}$ is $\emptyset$. (The number of solutions of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is zero.) $\square$

Solution (continued). The last matrix from the computation above is in reduced row echelon form, and we deduce that

$$\text{RREF}\Big(\left[\begin{array}{c:c} A & \mathbf{b} \end{array} \right] \Big) = \left[\begin{array}{cccc:c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

We see from $\text{RREF}(\left[\begin{array}{c:c} A & \mathbf{b} \end{array} \right])$ that the rightmost column of $\left[\begin{array}{c:c} A & \mathbf{b} \end{array} \right]$ is a pivot column; consequently, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is inconsistent, i.e. the solution set of the equation $A\mathbf{x} = \mathbf{b}$ is $\emptyset$. (The number of solutions of the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is zero.) $\square$

- Here is the row reduction again (next slide):

$$\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] = \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 1 & 0 & 1 & 0 & 2 \\ 2 & 2 & 1 & 1 & 0 \end{array}\right]$$

$$\underset{R_3 \to R_3 + R_1}{\overset{R_2 \to R_2 + 2R_1}{\sim}} \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 & 2 \end{array}\right]$$

$$\overset{R_3 \to R_3 + 2R_2}{\sim} \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

$$\overset{R_3 \to 2R_3}{\sim} \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\overset{R_1 \to R_1 + R_3}{\sim} \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\overset{R_1 \to R_1 + R_2}{\sim} \left[\begin{array}{cccc:c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

$$\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] \sim \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

$$\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

- We could in fact have stopped as soon as we got the red matrix (despite the fact that this matrix is not in reduced row echelon form).

$$\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] \sim \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

- We could in fact have stopped as soon as we got the red matrix (despite the fact that this matrix is not in reduced row echelon form).
- This is because the bottom row of the red matrix encodes the equation $0 = 2$, which is has no solutions.

$$\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right] \sim \left[\begin{array}{cccc:c} 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

- We could in fact have stopped as soon as we got the red matrix (despite the fact that this matrix is not in reduced row echelon form).
- This is because the bottom row of the red matrix encodes the equation $0 = 2$, which is has no solutions.
- Indeed, as soon as we obtain a row of the form $\left[\begin{array}{ccc:c} 0 & \dots & 0 & \blacksquare \end{array}\right]$, where $\blacksquare$ is a non-zero number, we can stop row reducing, and we can deduce that the system has no solutions (because this row encodes the equation $0 = \blacksquare$, and $\blacksquare$ is non-zero).

4. The rank of a matrix

- The *rank* of a matrix A (with entries in some field $\mathbb{F}$), denoted by $\operatorname{rank}(A)$, is the number of pivot columns of A.
- Equivalently, $\operatorname{rank}(A)$ is the number of pivot positions of A, or the number of non-zero rows of any row echelon form of A.
- To find the rank of a matrix, we first find some row echelon form of that matrix (e.g. by performing the forward phase of row reduction; the backward phase is optional), and we count the number of pivot columns (or alternatively, the number of pivot positions, or the number of non-zero rows) of that row echelon matrix.

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Example 1.6.1

Find the rank of each of the following matrices.

(a) $A = \begin{bmatrix} 0 & -3 & -6 & 3 & 4 & -1 \\ 2 & 1 & -4 & 13 & -4 & 3 \\ 2 & 3 & 0 & 11 & -6 & 5 \end{bmatrix}$, with entries understood to be in $\mathbb{R}$;

(b) $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}$, with entries understood to be in $\mathbb{Z}_3$.

Example 1.6.1

Find the rank of each of the following matrices.

(a) $A = \begin{bmatrix} 0 & -3 & -6 & 3 & 4 & -1 \\ 2 & 1 & -4 & 13 & -4 & 3 \\ 2 & 3 & 0 & 11 & -6 & 5 \end{bmatrix}$, with entries understood to be in $\mathbb{R}$;

Solution#1.

Example 1.6.1

Find the rank of each of the following matrices.

(a) $A = \begin{bmatrix} 0 & -3 & -6 & 3 & 4 & -1 \\ 2 & 1 & -4 & 13 & -4 & 3 \\ 2 & 3 & 0 & 11 & -6 & 5 \end{bmatrix}$, with entries understood to be in $\mathbb{R}$;

Solution#1. (a) In Example 1.3.9, we computed

$$\mathrm{RREF}(A) \quad = \quad \begin{bmatrix} 1 & 0 & -3 & 7 & 0 & 4 \\ 0 & 1 & 2 & -1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

Example 1.6.1

Find the rank of each of the following matrices.

(a) $A = \begin{bmatrix} 0 & -3 & -6 & 3 & 4 & -1 \\ 2 & 1 & -4 & 13 & -4 & 3 \\ 2 & 3 & 0 & 11 & -6 & 5 \end{bmatrix}$, with entries understood to be in $\mathbb{R}$;

Solution#1. (a) In Example 1.3.9, we computed

$$\text{RREF}(A) \;=\; \begin{bmatrix} 1 & 0 & -3 & 7 & 0 & 4 \\ 0 & 1 & 2 & -1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

The matrix $\text{RREF}(A)$ has three pivot columns (equivalently: three pivot positions or three non-zero rows), and so $\text{rank}(A) = 3$. $\square$

Example 1.6.1

Find the rank of each of the following matrices.

(a) $A = \begin{bmatrix} 0 & -3 & -6 & 3 & 4 & -1 \\ 2 & 1 & -4 & 13 & -4 & 3 \\ 2 & 3 & 0 & 11 & -6 & 5 \end{bmatrix}$, with entries understood to be in $\mathbb{R}$;

Solution#2.

Example 1.6.1

Find the rank of each of the following matrices.

(a) $A = \begin{bmatrix} 0 & -3 & -6 & 3 & 4 & -1 \\ 2 & 1 & -4 & 13 & -4 & 3 \\ 2 & 3 & 0 & 11 & -6 & 5 \end{bmatrix}$, with entries understood to be in $\mathbb{R}$;

Solution#2. (a) In Example 1.3.9, we saw that the matrix A is row equivalent to the following matrix in row echelon form:

$$\begin{bmatrix} 2 & 3 & 0 & 11 & -6 & 5 \\ 0 & -2 & -4 & 2 & 2 & -2 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

Example 1.6.1

Find the rank of each of the following matrices.

(a) $A = \begin{bmatrix} 0 & -3 & -6 & 3 & 4 & -1 \\ 2 & 1 & -4 & 13 & -4 & 3 \\ 2 & 3 & 0 & 11 & -6 & 5 \end{bmatrix}$, with entries understood to be in $\mathbb{R}$;

Solution#2. (a) In Example 1.3.9, we saw that the matrix A is row equivalent to the following matrix in row echelon form:

$$\begin{bmatrix} 2 & 3 & 0 & 11 & -6 & 5 \\ 0 & -2 & -4 & 2 & 2 & -2 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

This row echelon matrix has three pivot columns (equivalently: three pivot positions or three non-zero rows), and so $\text{rank}(A) = 3$. $\square$

Example 1.6.1

Find the rank of each of the following matrices.

(b) $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}$, with entries understood to be in $\mathbb{Z}_3$.

Solution#1.

Example 1.6.1

Find the rank of each of the following matrices.

(b) $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}$, with entries understood to be in $\mathbb{Z}_3$.

Solution#1. (b) In Example 1.3.10, we computed

$$\text{RREF}(B) = \begin{bmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 1.6.1

Find the rank of each of the following matrices.

- (b) $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}$, with entries understood to be in $\mathbb{Z}_3$.

Solution#1. (b) In Example 1.3.10, we computed

$$\text{RREF}(B) = \begin{bmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix $\text{RREF}(B)$ has three pivot columns (equivalently: three pivot positions or three non-zero rows), and so $\text{rank}(B) = 3$. $\square$

Example 1.6.1

Find the rank of each of the following matrices.

(b) $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}$, with entries understood to be in $\mathbb{Z}_3$.

Solution#2.

Example 1.6.1

Find the rank of each of the following matrices.

(b) $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}$, with entries understood to be in $\mathbb{Z}_3$.

Solution#2. (b) In Example 1.3.10, we saw that the matrix B is row equivalent to the following matrix in row echelon form:

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

Example 1.6.1

Find the rank of each of the following matrices.

(b) $B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}$, with entries understood to be in $\mathbb{Z}_3$.

Solution#2. (b) In Example 1.3.10, we saw that the matrix B is row equivalent to the following matrix in row echelon form:

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

This row echelon matrix has three pivot columns (equivalently: three pivot positions or three non-zero rows), and so $\text{rank}(B) = 3$. $\square$

Corollary 1.3.8

Two matrices (with entries in some field) are row equivalent iff they have the same reduced row echelon form.

Corollary 1.3.8

Two matrices (with entries in some field) are row equivalent iff they have the same reduced row echelon form.

Proposition 1.6.2

Row equivalent matrices (with entries in some field) have the same rank.

Proof (outline).

Corollary 1.3.8

Two matrices (with entries in some field) are row equivalent iff they have the same reduced row echelon form.

Proposition 1.6.2

Row equivalent matrices (with entries in some field) have the same rank.

Proof (outline). By Corollary 1.3.8, row equivalent matrices have te same RREF.

Corollary 1.3.8

Two matrices (with entries in some field) are row equivalent iff they have the same reduced row echelon form.

Proposition 1.6.2

Row equivalent matrices (with entries in some field) have the same rank.

Proof (outline). By Corollary 1.3.8, row equivalent matrices have te same RREF. By definition, the rank of a matrix can be read off from its RREF (it is simply the number of non-zero rows of the RREF). $\square$

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposition 1.6.3

Let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.[a]

[a] This means that $\text{rank}(A) \leq n$ (i.e. $\text{rank}(A)$ is at most the number of rows of A) and $\text{rank}(A) \leq m$ (i.e. $\text{rank}(A)$ is at most the number of columns of A).

- Proof: Follows immediately from the definition of rank. (The full details are in the lecture notes.)

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.[a]

---

[a] This means that $\text{rank}(A) \leq n$ (i.e. $\text{rank}(A)$ is at most the number of rows of A) and $\text{rank}(A) \leq m$ (i.e. $\text{rank}(A)$ is at most the number of columns of A).

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.[a]

[a] This means that $\text{rank}(A) \leq n$ (i.e. $\text{rank}(A)$ is at most the number of rows of A) and $\text{rank}(A) \leq m$ (i.e. $\text{rank}(A)$ is at most the number of columns of A).

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.[a]

---

[a] This means that $\text{rank}(A) \leq n$ (i.e. $\text{rank}(A)$ is at most the number of rows of A) and $\text{rank}(A) \leq m$ (i.e. $\text{rank}(A)$ is at most the number of columns of A).

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
  - if $\text{rank}(A) = n$, then A is said to have *full row rank*;

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.[a]

[a] This means that $\text{rank}(A) \leq n$ (i.e. $\text{rank}(A)$ is at most the number of rows of A) and $\text{rank}(A) \leq m$ (i.e. $\text{rank}(A)$ is at most the number of columns of A).

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
  - if $\text{rank}(A) = n$, then A is said to have *full row rank*;
  - if $\text{rank}(A) = m$, then A is said to have *full column rank*;

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.[a]

---

[a] This means that $\text{rank}(A) \leq n$ (i.e. $\text{rank}(A)$ is at most the number of rows of A) and $\text{rank}(A) \leq m$ (i.e. $\text{rank}(A)$ is at most the number of columns of A).

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
  - if $\text{rank}(A) = n$, then A is said to have *full row rank*;
  - if $\text{rank}(A) = m$, then A is said to have *full column rank*;
  - if $\text{rank}(A) = \min\{n, m\}$, then A is said to have *full rank*;

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.[a]

[a] This means that $\text{rank}(A) \leq n$ (i.e. $\text{rank}(A)$ is at most the number of rows of A) and $\text{rank}(A) \leq m$ (i.e. $\text{rank}(A)$ is at most the number of columns of A).

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
  - if $\text{rank}(A) = n$, then A is said to have *full row rank*;
  - if $\text{rank}(A) = m$, then A is said to have *full column rank*;
  - if $\text{rank}(A) = \min\{n, m\}$, then A is said to have *full rank*;
  - if $\text{rank}(A) < \min\{n, m\}$, then A is said to be *rank-deficient*.

- As Theorem 1.6.4 (next slide) shows, the number of solutions of a matrix-vector equation $A\mathbf{x} = \mathbf{b}$ can easily be determined if we know the size of the matrix A (i.e. the number of rows and columns of A) and we also know rank(A) and rank($\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]$).
  - The detailed proof of Theorem 1.6.4 is in the Lecture Notes.
  - An outline of the proof: on the board!

Theorem 1.6.4

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times m}$ and $\mathbf{b} \in \mathbb{F}^n$. Then

$$\text{rank}(A) \quad \leq \quad \text{rank}(\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]) \quad \leq \quad \text{rank}(A) + 1.$$

Moreover, all the following hold:

- (a) if $\text{rank}(\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]) \neq \text{rank}(A)$ (and consequently, $\text{rank}(\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]) = \text{rank}(A) + 1$), then $A\mathbf{x} = \mathbf{b}$ is inconsistent.
- (b) if $\text{rank}(\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]) = \text{rank}(A) = m$, then $A\mathbf{x} = \mathbf{b}$ has a unique solution.
- (c) if $\text{rank}(\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]) = \text{rank}(A) < m$, then $A\mathbf{x} = \mathbf{b}$ has more than one solution, and more precisely,
  - (c.1) if the field $\mathbb{F}$ is finite, then $A\mathbf{x} = \mathbf{b}$ has exactly $|\mathbb{F}|^{m-\text{rank}(A)}$ many solutions,
  - (c.2) if the field $\mathbb{F}$ is infinite, then $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions.

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
  - if $\text{rank}(A) = n$, then A is said to have *full row rank*;
  - if $\text{rank}(A) = m$, then A is said to have *full column rank*;
  - if $\text{rank}(A) = \min\{n, m\}$, then A is said to have *full rank*;
  - if $\text{rank}(A) < \min\{n, m\}$, then A is said to be *rank-deficient*.

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
  - if $\text{rank}(A) = n$, then A is said to have *full row rank*;
  - if $\text{rank}(A) = m$, then A is said to have *full column rank*;
  - if $\text{rank}(A) = \min\{n, m\}$, then A is said to have *full rank*;
  - if $\text{rank}(A) < \min\{n, m\}$, then A is said to be *rank-deficient*.
- Our next goal is to prove a derive a couple of corollaries of Theorem 1.6.4 for matrices of full rank.

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
    - if $\text{rank}(A) = n$, then A is said to have *full row rank*;
    - if $\text{rank}(A) = m$, then A is said to have *full column rank*;
    - if $\text{rank}(A) = \min\{n, m\}$, then A is said to have *full rank*;
    - if $\text{rank}(A) < \min\{n, m\}$, then A is said to be *rank-deficient*.
- Our next goal is to prove a derive a couple of corollaries of Theorem 1.6.4 for matrices of full rank.
- By definition, a matrix of full rank has full column rank or full row rank (possibly both). We deal with these two cases separately.

Proposition 1.6.3

Let $\mathbb{F}$ be a field, and let A be an $n \times m$ matrix (with entries in some field $\mathbb{F}$). Then $\text{rank}(A) \leq \min\{n, m\}$.

- **Terminology:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times m}$ (so, A has n rows and m columns):
    - if $\text{rank}(A) = n$, then A is said to have *full row rank*;
    - if $\text{rank}(A) = m$, then A is said to have *full column rank*;
    - if $\text{rank}(A) = \min\{n, m\}$, then A is said to have *full rank*;
    - if $\text{rank}(A) < \min\{n, m\}$, then A is said to be *rank-deficient*.

- Our next goal is to prove a derive a couple of corollaries of Theorem 1.6.4 for matrices of full rank.
- By definition, a matrix of full rank has full column rank or full row rank (possibly both). We deal with these two cases separately.
- After that, we prove Theorem 1.6.8, which deals with **square** matrices of full rank.
    - Note that such matrices have both full column rank and full row rank.

- In a matrix of full column rank, all columns are pivot columns.

- In a matrix of full column rank, all columns are pivot columns.
- So, the reduced row echelon form of such a matrix is of the form

$$\left[\begin{array}{cccccc} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ \hdashline 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \end{array}\right].$$

- In a matrix of full column rank, all columns are pivot columns.
- So, the reduced row echelon form of such a matrix is of the form

$$\left[\begin{array}{cccccc} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ \hdashline 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \end{array}\right].$$

- More precisely, if we have an $n \times m$ matrix of full column rank, then the reduced row echelon form of that matrix is obtained from the identity matrix I_m by adding $n - m$ many zero rows to the bottom.

- A *homogeneous matrix-vector equation* is a matrix-vector equation of the form $A\mathbf{x} = \mathbf{0}$.

- A *homogeneous matrix-vector equation* is a matrix-vector equation of the form $A\mathbf{x} = \mathbf{0}$.
- Note that such an equation is always consistent: indeed, $\mathbf{x} = \mathbf{0}$ is a solution, called the trivial solution.

- A *homogeneous matrix-vector equation* is a matrix-vector equation of the form $A\mathbf{x} = \mathbf{0}$.
- Note that such an equation is always consistent: indeed, $\mathbf{x} = \mathbf{0}$ is a solution, called the trivial solution.

Corollary 1.6.5

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

(a) $\text{rank}(A) = m$ (i.e. A has full column rank);
(b) the homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution (i.e. the solution $\mathbf{x} = \mathbf{0}$);
(c) there exists some vector $\mathbf{b} \in \mathbb{F}^n$ s.t. the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;
(d) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has at most one solution.

Corollary 1.6.5

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

(a) $\text{rank}(A) = m$ (i.e. A has full column rank);

(b) the homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution (i.e. the solution $\mathbf{x} = \mathbf{0}$);

(c) there exists some vector $\mathbf{b} \in \mathbb{F}^n$ s.t. the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;

(d) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has at most one solution.

Proof (outline).

Corollary 1.6.5

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

- (a) $\text{rank}(A) = m$ (i.e. A has full column rank);
- (b) the homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution (i.e. the solution $\mathbf{x} = \mathbf{0}$);
- (c) there exists some vector $\mathbf{b} \in \mathbb{F}^n$ s.t. the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;
- (d) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has at most one solution.

Proof (outline). It is enough to prove following implications:

$$\begin{array}{ccc} (a) & \Longrightarrow & (d) \\ \Uparrow & & \Downarrow \\ (c) & \Longleftarrow & (b) \end{array}$$

Corollary 1.6.5

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

(a) $\text{rank}(A) = m$ (i.e. A has full column rank);

(b) the homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution (i.e. the solution $\mathbf{x} = \mathbf{0}$);

(c) there exists some vector $\mathbf{b} \in \mathbb{F}^n$ s.t. the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;

(d) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has at most one solution.

Proof (outline). It is enough to prove following implications:

$$\begin{array}{ccc} \text{(a)} & \Longrightarrow & \text{(d)} \\ \Uparrow & & \Downarrow \\ \text{(c)} & \Longleftarrow & \text{(b)} \end{array}$$

The implications "(d) $\Longrightarrow$ (b)" and "(b) $\Longrightarrow$ (c)" are obvious,

Corollary 1.6.5

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

(a) $\text{rank}(A) = m$ (i.e. A has full column rank);

(b) the homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution (i.e. the solution $\mathbf{x} = \mathbf{0}$);

(c) there exists some vector $\mathbf{b} \in \mathbb{F}^n$ s.t. the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;

(d) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has at most one solution.

Proof (outline). It is enough to prove following implications:

$$\begin{array}{ccc} (a) & \Longrightarrow & (d) \\ \Uparrow & & \Downarrow \\ (c) & \Longleftarrow & (b) \end{array}$$

The implications "(d) $\Longrightarrow$ (b)" and "(b) $\Longrightarrow$ (c)" are obvious, and "(c) $\Longrightarrow$ (a)" and "(a) $\Longrightarrow$ (d) follow from Theorem 1.6.4." $\square$

- We now consider matrices of full row rank.

- We now consider matrices of full row rank.
- Note that matrices of full row rank are precisely those matrices whose reduced row echelon form has no zero rows.

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & * & 0 & 0 & * & * & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & * & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \end{bmatrix}$$

- We now consider matrices of full row rank.
- Note that matrices of full row rank are precisely those matrices whose reduced row echelon form has no zero rows.

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & * & 0 & 0 & * & * & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & * & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \end{bmatrix}$$

Corollary 1.6.6

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

- (a) $\text{rank}(A) = n$ (i.e. A has full row rank);
- (b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof. Suppose first that (a) holds. We must prove (b).

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof. Suppose first that (a) holds. We must prove (b). Fix any $\mathbf{b} \in \mathbb{F}^n$.

Corollary 1.6.6

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

- (a) $\text{rank}(A) = n$ (i.e. A has full row rank);
- (b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof. Suppose first that (a) holds. We must prove (b). Fix any $\mathbf{b} \in \mathbb{F}^n$. Then

$$\begin{array}{rcll} n & = & \text{rank}(A) & \text{by (a)} \\ & \leq & \text{rank}(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]) & \text{by Theorem 1.6.4} \\ & \leq & n & \text{by Proposition 1.6.3,} \end{array}$$

and it follows that $\text{rank}(\left[\begin{array}{c:c} A & \mathbf{b} \end{array}\right]) = \text{rank}(A) = n$. But now Theorem 1.6.4 guarantees that the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent. Thus, (b) holds.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued).

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued). Suppose now that (a) is false; we must show that (b) is false, i.e. that there exists some $\mathbf{b} \in \mathbb{F}^n$ s.t. $A\mathbf{x} = \mathbf{b}$ is inconsistent.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued). Suppose now that (a) is false; we must show that (b) is false, i.e. that there exists some $\mathbf{b} \in \mathbb{F}^n$ s.t. $A\mathbf{x} = \mathbf{b}$ is inconsistent. Since A is an $n \times m$ matrix and $\text{rank}(A) \neq n$, Proposition 1.6.3 guarantees that $\text{rank}(A) \leq n - 1$. Now, set $U := \text{RREF}(A)$, and let $R_1, \dots, R_k$ be some sequence of elementary row operations that transforms A into U, and for each $i \in \{1, \dots, k\}$, let R_i' be the elementary row operation that reverses (undoes) the elementary row operation R_i.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued). Suppose now that (a) is false; we must show that (b) is false, i.e. that there exists some $\mathbf{b} \in \mathbb{F}^n$ s.t. $A\mathbf{x} = \mathbf{b}$ is inconsistent. Since A is an $n \times m$ matrix and $\text{rank}(A) \neq n$, Proposition 1.6.3 guarantees that $\text{rank}(A) \leq n - 1$. Now, set $U := \text{RREF}(A)$, and let $R_1, \dots, R_k$ be some sequence of elementary row operations that transforms A into U, and for each $i \in \{1, \dots, k\}$, let R_i' be the elementary row operation that reverses (undoes) the elementary row operation R_i. Since U has n rows and $r := \text{rank}(A) \leq n - 1$, we see that the $(r + 1)$-th row of U is a zero row.

Corollary 1.6.6

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

- (a) $\text{rank}(A) = n$ (i.e. A has full row rank);
- (b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued). Suppose now that (a) is false; we must show that (b) is false, i.e. that there exists some $\mathbf{b} \in \mathbb{F}^n$ s.t. $A\mathbf{x} = \mathbf{b}$ is inconsistent. Since A is an $n \times m$ matrix and $\text{rank}(A) \neq n$, Proposition 1.6.3 guarantees that $\text{rank}(A) \leq n - 1$. Now, set $U := \text{RREF}(A)$, and let $R_1, \dots, R_k$ be some sequence of elementary row operations that transforms A into U, and for each $i \in \{1, \dots, k\}$, let R_i' be the elementary row operation that reverses (undoes) the elementary row operation R_i. Since U has n rows and $r := \text{rank}(A) \leq n - 1$, we see that the $(r + 1)$-th row of U is a zero row. Then the rightmost column of the matrix $\left[\begin{array}{c|c} U & \mathbf{e}_{r+1} \end{array}\right]$ is a pivot column, and consequently, $U\mathbf{x} = \mathbf{e}_{r+1}$ is inconsistent.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued). Now, we perform the elementary row operations $R'_k, \dots, R'_1$ on the matrix $\left[\begin{array}{c|c} U & \mathbf{e}_{r+1} \end{array}\right]$, and we obtain the matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ for some vector $\mathbf{b} \in \mathbb{F}^n$.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued). Now, we perform the elementary row operations $R_k', \dots, R_1'$ on the matrix $\left[\begin{array}{c|c} U & \mathbf{e}_{r+1} \end{array}\right]$, and we obtain the matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ for some vector $\mathbf{b} \in \mathbb{F}^n$. Since matrices $\left[\begin{array}{c|c} U & \mathbf{e}_{r+1} \end{array}\right]$ and $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ are row equivalent, the matrix-vector equations $U\mathbf{x} = \mathbf{e}_{r+1}$ and $A\mathbf{x} = \mathbf{b}$ are equivalent.

Corollary 1.6.6

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

(a) $\text{rank}(A) = n$ (i.e. A has full row rank);

(b) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

Proof (continued). Now, we perform the elementary row operations $R'_k, \dots, R'_1$ on the matrix $\left[\begin{array}{c|c} U & \mathbf{e}_{r+1} \end{array}\right]$, and we obtain the matrix $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ for some vector $\mathbf{b} \in \mathbb{F}^n$. Since matrices $\left[\begin{array}{c|c} U & \mathbf{e}_{r+1} \end{array}\right]$ and $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$ are row equivalent, the matrix-vector equations $U\mathbf{x} = \mathbf{e}_{r+1}$ and $A\mathbf{x} = \mathbf{b}$ are equivalent. Since the matrix-vector equation $U\mathbf{x} = \mathbf{e}_{r+1}$ is inconsistent, it follows that the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is also inconsistent. Thus, (b) is false. $\square$

- We now consider the special case of square matrices of full rank.

- We now consider the special case of square matrices of full rank.

Proposition 1.6.7

Let $\mathbb{F}$ be a field. Then for all square matrices $A \in \mathbb{F}^{n \times n}$, we have that $\text{rank}(A) = n$ iff $\text{RREF}(A) = I_n$. In particular, $\text{rank}(I_n) = n$.

Proof.

- We now consider the special case of square matrices of full rank.

Proposition 1.6.7

Let $\mathbb{F}$ be a field. Then for all square matrices $A \in \mathbb{F}^{n \times n}$, we have that $\text{rank}(A) = n$ iff $\text{RREF}(A) = I_n$. In particular, $\text{rank}(I_n) = n$.

Proof. I_n is a matrix in reduced row echelon form, and it has n pivot columns; so, $\text{rank}(I_n) = n$. Moreover, it is clear that I_n is the **only** reduced row echelon form matrix in $\mathbb{F}^{n \times n}$ of rank n.

Now, fix any matrix $A \in \mathbb{F}^{n \times n}$.

- We now consider the special case of square matrices of full rank.

Proposition 1.6.7

Let $\mathbb{F}$ be a field. Then for all square matrices $A \in \mathbb{F}^{n \times n}$, we have that $\text{rank}(A) = n$ iff $\text{RREF}(A) = I_n$. In particular, $\text{rank}(I_n) = n$.

Proof. I_n is a matrix in reduced row echelon form, and it has n pivot columns; so, $\text{rank}(I_n) = n$. Moreover, it is clear that I_n is the **only** reduced row echelon form matrix in $\mathbb{F}^{n \times n}$ of rank n.

Now, fix any matrix $A \in \mathbb{F}^{n \times n}$. By Proposition 1.6.2, we have that $\text{rank}(A) = \text{rank}\big(\text{RREF}(A)\big)$. Since I_n is the only reduced row echelon form matrix in $\mathbb{F}^{n \times n}$ of rank n, it follows that $\text{rank}(A) = n$ iff $\text{RREF}(A) = I_n$. $\square$

Theorem 1.6.8

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a **square** matrix. Then the following are equivalent:

(a) $\text{rank}(A) = n$ (i.e. the square matrix A has full rank);

(b) $\text{RREF}(A) = I_n$;

(c) the homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution (i.e. the solution $\mathbf{x} = \mathbf{0}$);

(d) there exists some vector $\mathbf{b} \in \mathbb{F}^n$ s.t. the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;

(e) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;

(f) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector $A\mathbf{x} = \mathbf{b}$ equation has at most one solution;

(g) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent.

- Proof: Lecture Notes.

5. Matrix operations

5. Matrix operations

- Suppose that $\mathbb{F}$ is a field.

5. Matrix operations

- Suppose that $\mathbb{F}$ is a field.
- Given matrices $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times m}$ and $B = \left[\begin{array}{c} b_{i,j} \end{array}\right]_{n\times m}$ in $\mathbb{F}^{n\times m}$, and given a scalar c, we define
    - $A + B := \left[\begin{array}{c} a_{i,j} + b_{i,j} \end{array}\right]_{n\times m}$;
    - $A - B := \left[\begin{array}{c} a_{i,j} - b_{i,j} \end{array}\right]_{n\times m}$;
    - $cA := \left[\begin{array}{c} ca_{i,j} \end{array}\right]$.

5. Matrix operations

- Suppose that $\mathbb{F}$ is a field.
- Given matrices $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times m}$ and $B = \left[\begin{array}{c} b_{i,j} \end{array}\right]_{n\times m}$ in $\mathbb{F}^{n\times m}$, and given a scalar c, we define
  - $A + B := \left[\begin{array}{c} a_{i,j} + b_{i,j} \end{array}\right]_{n\times m}$;
  - $A - B := \left[\begin{array}{c} a_{i,j} - b_{i,j} \end{array}\right]_{n\times m}$;
  - $cA := \left[\begin{array}{c} ca_{i,j} \end{array}\right]$.
- Thus, we add (resp. subtract) matrices by adding (resp. subtracting) corresponding entries, i.e.
  - $\left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times m} + \left[\begin{array}{c} b_{i,j} \end{array}\right]_{n\times m} = \left[\begin{array}{c} a_{i,j} + b_{i,j} \end{array}\right]_{n\times m}$;
  - $\left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times m} - \left[\begin{array}{c} b_{i,j} \end{array}\right]_{n\times m} = \left[\begin{array}{c} a_{i,j} - b_{i,j} \end{array}\right]_{n\times m}$.

5. Matrix operations

- Suppose that $\mathbb{F}$ is a field.
- Given matrices $A = \left[\ a_{i,j}\ \right]_{n\times m}$ and $B = \left[\ b_{i,j}\ \right]_{n\times m}$ in $\mathbb{F}^{n\times m}$, and given a scalar c, we define
  - $A + B := \left[\ a_{i,j} + b_{i,j}\ \right]_{n\times m}$;
  - $A - B := \left[\ a_{i,j} - b_{i,j}\ \right]_{n\times m}$;
  - $cA := \left[\ ca_{i,j}\ \right]$.
- Thus, we add (resp. subtract) matrices by adding (resp. subtracting) corresponding entries, i.e.
  - $\left[\ a_{i,j}\ \right]_{n\times m} + \left[\ b_{i,j}\ \right]_{n\times m} = \left[\ a_{i,j} + b_{i,j}\ \right]_{n\times m}$;
  - $\left[\ a_{i,j}\ \right]_{n\times m} - \left[\ b_{i,j}\ \right]_{n\times m} = \left[\ a_{i,j} - b_{i,j}\ \right]_{n\times m}$.
- Similarly, we multiply a matrix by a scalar (on the left) by multiplying each entry of the matrix by that scalar, i.e.
  - $c\left[\ a_{i,j}\ \right]_{n\times m} = \left[\ ca_{i,j}\ \right]_{n\times m}$.

- We can also multiply matrices!

- We can also multiply matrices!
- Let $\mathbb{F}$ be a field, and suppose that we are given two matrices, $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, where $B = \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_p \end{array} \right]$.

- We can also multiply matrices!
- Let $\mathbb{F}$ be a field, and suppose that we are given two matrices, $A \in \mathbb{F}^{n \times m}$ and $B \in \mathbb{F}^{m \times p}$, where $B = \left[\begin{array}{ccc} \mathbf{b}_1 & \dots & \mathbf{b}_p \end{array} \right]$.
- We define

$$AB \quad := \quad \left[\begin{array}{ccc} A\mathbf{b}_1 & \dots & A\mathbf{b}_p \end{array} \right]$$

Note that $AB \in \mathbb{F}^{n \times p}$.

- We can also multiply matrices!
- Let $\mathbb{F}$ be a field, and suppose that we are given two matrices, $A \in \mathbb{F}^{n \times m}$ and $B \in \mathbb{F}^{m \times p}$, where $B = \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_p \end{bmatrix}$.
- We define
$$AB \; := \; \begin{bmatrix} A\mathbf{b}_1 & \dots & A\mathbf{b}_p \end{bmatrix}$$
Note that $AB \in \mathbb{F}^{n \times p}$.
- Note that, for the product AB to be defined, the number of **columns** of A must be the same as the number of **rows** of B.

- We can also multiply matrices!
- Let $\mathbb{F}$ be a field, and suppose that we are given two matrices, $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, where $B = \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_p \end{bmatrix}$.
- We define
$$AB \quad := \quad \begin{bmatrix} A\mathbf{b}_1 & \dots & A\mathbf{b}_p \end{bmatrix}$$
  Note that $AB \in \mathbb{F}^{n\times p}$.
- Note that, for the product AB to be defined, the number of **columns** of A must be the same as the number of **rows** of B.
- The matrix AB has the same number of **rows** as A, and the same number of **columns** as B.

- We can also multiply matrices!
- Let $\mathbb{F}$ be a field, and suppose that we are given two matrices, $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, where $B = \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_p \end{bmatrix}$.
- We define
$$AB \;:=\; \begin{bmatrix} A\mathbf{b}_1 & \dots & A\mathbf{b}_p \end{bmatrix}$$
Note that $AB \in \mathbb{F}^{n\times p}$.
- Note that, for the product AB to be defined, the number of **columns** of A must be the same as the number of **rows** of B.
- The matrix AB has the same number of **rows** as A, and the same number of **columns** as B.
- Schematically, we get:
$$(n \times m) \;\cdot\; (m \times p) \;=\; (n \times p).$$

Example 1.7.1

Let

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ -2 & 1 \\ 0 & -1 \end{bmatrix},$$

with entries understood to be in $\mathbb{R}$. Compute AB.

Solution.

Example 1.7.1

Let

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ -2 & 1 \\ 0 & -1 \end{bmatrix},$$

with entries understood to be in $\mathbb{R}$. Compute AB.

Solution. We set

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix},$$

so that $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix}$.

Example 1.7.1

Let

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ -2 & 1 \\ 0 & -1 \end{bmatrix},$$

with entries understood to be in $\mathbb{R}$. Compute AB.

Solution. We set

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix},$$

so that $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix}$. Then $AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix}$.

Example 1.7.1

Let

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ -2 & 1 \\ 0 & -1 \end{bmatrix},$$

with entries understood to be in $\mathbb{R}$. Compute AB.

Solution. We set

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix},$$

so that $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix}$. Then $AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix}$. We compute $A\mathbf{b}_1 = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$ and $A\mathbf{b}_2 = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$, which yields $AB = \begin{bmatrix} -3 & 4 \\ 6 & -4 \end{bmatrix}$. $\square$

Proposition 1.7.2

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

- (a) $I_n A = A I_m = A$;
- (b) $A O_{m \times p} = O_{n \times p}$;
- (c) $O_{p \times n} A = O_{p \times m}$.

Proof.

Proposition 1.7.2

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

a. $I_n A = A I_m = A$;
b. $A O_{m \times p} = O_{n \times p}$;
c. $O_{p \times n} A = O_{p \times m}$.

Proof. Parts (b) and (c) readily follow from the appropriate definitions (the details are left as an easy exercise). Let us prove (a).

Proposition 1.7.2

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

(a) $I_n A = A I_m = A$;

(b) $A O_{m \times p} = O_{n \times p}$;

(c) $O_{p \times n} A = O_{p \times m}$.

Proof. Parts (b) and (c) readily follow from the appropriate definitions (the details are left as an easy exercise). Let us prove (a). Set $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array} \right]$. To show that $I_n A = A$, we compute (next slide):

Proposition 1.7.2

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

(a) $I_n A = A I_m = A$;

Proof (continued). Reminder: $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

$$\begin{aligned} I_n A &= I_n \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \\ &= \begin{bmatrix} I_n\mathbf{a}_1 & \dots & I_n\mathbf{a}_m \end{bmatrix} && \text{by the definition of matrix multiplication} \\ &= \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} && \text{by Proposition 1.4.5} \\ &= A. \end{aligned}$$

Proposition 1.7.2

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

(a) $I_n A = A I_m = A$;

Proof (continued). Reminder: $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

Proposition 1.7.2

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

(a) $I_n A = A I_m = A$;

Proof (continued). Reminder: $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array}\right]$.
On the other hand, to show that $AI_m = A$, we compute:

$$\begin{array}{rcll} AI_m & = & A\left[\begin{array}{ccc} \mathbf{e}_1^m & \dots & \mathbf{e}_m^m \end{array}\right] & \\ & = & \left[\begin{array}{ccc} A\mathbf{e}_1^m & \dots & A\mathbf{e}_m^m \end{array}\right] & \text{by the definition of matrix multiplication} \\ & = & \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{array}\right] & \text{by Proposition 1.4.4} \\ & = & A. & \end{array}$$

This proves (a). □

Proposition 1.7.2

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

(a) $I_n A = A I_m = A$;

(b) $A O_{m \times p} = O_{n \times p}$;

(c) $O_{p \times n} A = O_{p \times m}$.

- There is another way to compute the product of two matrices.

- There is another way to compute the product of two matrices.
- Suppose we are given matrices $A = \left[\ a_{i,j}\ \right]_{n\times m}$ and $B = \left[\ b_{i,j}\ \right]_{m\times p}$, both with entries in some field $\mathbb{F}$.

- There is another way to compute the product of two matrices.
- Suppose we are given matrices $A = \left[\ a_{i,j}\ \right]_{n\times m}$ and $B = \left[\ b_{i,j}\ \right]_{m\times p}$, both with entries in some field $\mathbb{F}$.
- Then AB is an $n \times p$ matrix whose i,j-th entry is

$$\sum_{k=1}^{m} a_{i,k} b_{k,j},$$

  for all indices $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, p\}$.
    - Justification: Lecture Notes (not very hard, but notationally a bit messy).

- There is another way to compute the product of two matrices.
- Suppose we are given matrices $A = \left[\ a_{i,j}\ \right]_{n \times m}$ and $B = \left[\ b_{i,j}\ \right]_{m \times p}$, both with entries in some field $\mathbb{F}$.
- Then AB is an $n \times p$ matrix whose i, j-th entry is

$$\sum_{k=1}^{m} a_{i,k} b_{k,j},$$

for all indices $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, p\}$.
  - Justification: Lecture Notes (not very hard, but notationally a bit messy).
- Another way to write this is as follows:

$$\left[\ a_{i,j}\ \right]_{n \times m} \left[\ b_{i,j}\ \right]_{m \times p} = \left[\ \sum_{k=1}^{m} a_{i,k} b_{k,j}\ \right]_{n \times p},$$

where in each of the three matrices, the expression between the square brackets is the general form of the i, j-th entry (i.e. the entry in the i-th row and j-th column) of the matrix in question.

- To obtain the i,j-th entry of the matrix AB, we focus on the i-th row of A and j-th column of B.
- We then take the sum of the products of the corresponding entries of this row and column, and we obtain the i,j-th entry of AB.

$$
\begin{array}{cc}
 & \begin{bmatrix} b_{1,1} & \dots & b_{1,j} & \dots & b_{1,p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{k,1} & \dots & b_{k,j} & \dots & b_{k,p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{m,1} & \dots & b_{m,j} & \dots & b_{m,p} \end{bmatrix} \\
\begin{bmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,m} \end{bmatrix} & \begin{bmatrix} \\ \\ \displaystyle\sum_{k=1}^{m} a_{i,k} b_{k,j} \\ \\ \\ \end{bmatrix}
\end{array}
$$

Example 1.7.3

Let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_2$. Compute the matrix AB.

Solution.

Example 1.7.3

Let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_2$. Compute the matrix AB.

Solution. We compute as shown below (the rows of A are color coded, as are the columns of B).

$$\begin{array}{cc} & \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 1\cdot 1 + 0 \cdot 1 & 1 \cdot 0 + 0 \cdot 1 & 1 \cdot 1 + 0 \cdot 0 \\ 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 0 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot 0 \end{bmatrix} \end{array}$$

Example 1.7.3

Let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_2$. Compute the matrix AB.

Solution (continued).

$$AB = \begin{bmatrix} 1 \cdot 1 + 0 \cdot 1 & 1 \cdot 0 + 0 \cdot 1 & 1 \cdot 1 + 0 \cdot 0 \\ 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 0 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot 0 \end{bmatrix}$$

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

$\square$

Theorem 1.7.5

For any matrices A, B, and C, and any scalars α and β, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalars all belong to the same field $\mathbb{F}$):

a. $(\alpha + \beta)A = \alpha A + \beta A$;
b. $(\alpha\beta)A = \alpha(\beta A)$
c. $A + B = B + A$;
d. $(A + B) + C = A + (B + C)$;
e. $(A + B)C = AC + BC$;
f. $A(B + C) = AB + AC$;
g. $(AB)C = A(BC)$;
h. $(\alpha A)B = \alpha(AB)$;
i. $A(\alpha B) = \alpha(AB)$.

- The only difficult part of Theorem 1.7.5 is (g).
  - So, let us prove that.

- The only difficult part of Theorem 1.7.5 is (g).
  - So, let us prove that.

(g) $(AB)C = A(BC)$

Proof of (g).

- The only difficult part of Theorem 1.7.5 is (g).
  - So, let us prove that.

(g) $(AB)C = A(BC)$

Proof of (g). Fix matrices $A = \left[\ a_{i,j}\ \right]_{n_1 \times n_2}$ in $\mathbb{F}^{n_1 \times n_2}$, $B = \left[\ b_{i,j}\ \right]_{n_2 \times n_3}$ in $\mathbb{F}^{n_2 \times n_3}$, and $C = \left[\ c_{i,j}\ \right]_{n_3 \times n_4}$ in $\mathbb{F}^{n_3 \times n_4}$.

- The only difficult part of Theorem 1.7.5 is (g).
  - So, let us prove that.

(g) $(AB)C = A(BC)$

Proof of (g). Fix matrices $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n_1 \times n_2}$ in $\mathbb{F}^{n_1 \times n_2}$, $B = \begin{bmatrix} b_{i,j} \end{bmatrix}_{n_2 \times n_3}$ in $\mathbb{F}^{n_2 \times n_3}$, and $C = \begin{bmatrix} c_{i,j} \end{bmatrix}_{n_3 \times n_4}$ in $\mathbb{F}^{n_3 \times n_4}$.

Clearly, both $(AB)C$ and $A(BC)$ are matrices in $\mathbb{F}^{n_1 \times n_4}$.

- The only difficult part of Theorem 1.7.5 is (g).
    - So, let us prove that.

(g) $(AB)C = A(BC)$

Proof of (g). Fix matrices $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n_1 \times n_2}$ in $\mathbb{F}^{n_1 \times n_2}$, $B = \left[\begin{array}{c} b_{i,j} \end{array}\right]_{n_2 \times n_3}$ in $\mathbb{F}^{n_2 \times n_3}$, and $C = \left[\begin{array}{c} c_{i,j} \end{array}\right]_{n_3 \times n_4}$ in $\mathbb{F}^{n_3 \times n_4}$.

Clearly, both $(AB)C$ and $A(BC)$ are matrices in $\mathbb{F}^{n_1 \times n_4}$. To prove that these two matrices are equal, it suffices to prove that their corresponding entries are equal.

- The only difficult part of Theorem 1.7.5 is (g).
    - So, let us prove that.

(g) $(AB)C = A(BC)$

Proof of (g). Fix matrices $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n_1 \times n_2}$ in $\mathbb{F}^{n_1 \times n_2}$, $B = \left[\begin{array}{c} b_{i,j} \end{array}\right]_{n_2 \times n_3}$ in $\mathbb{F}^{n_2 \times n_3}$, and $C = \left[\begin{array}{c} c_{i,j} \end{array}\right]_{n_3 \times n_4}$ in $\mathbb{F}^{n_3 \times n_4}$.

Clearly, both $(AB)C$ and $A(BC)$ are matrices in $\mathbb{F}^{n_1 \times n_4}$. To prove that these two matrices are equal, it suffices to prove that their corresponding entries are equal. So, fix indices $i \in \{1, \dots, n_1\}$ and $j \in \{1, \dots, n_4\}$.

- The only difficult part of Theorem 1.7.5 is (g).
    - So, let us prove that.

(g) $(AB)C = A(BC)$

Proof of (g). Fix matrices $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n_1 \times n_2}$ in $\mathbb{F}^{n_1 \times n_2}$, $B = \left[\begin{array}{c} b_{i,j} \end{array}\right]_{n_2 \times n_3}$ in $\mathbb{F}^{n_2 \times n_3}$, and $C = \left[\begin{array}{c} c_{i,j} \end{array}\right]_{n_3 \times n_4}$ in $\mathbb{F}^{n_3 \times n_4}$.

Clearly, both $(AB)C$ and $A(BC)$ are matrices in $\mathbb{F}^{n_1 \times n_4}$. To prove that these two matrices are equal, it suffices to prove that their corresponding entries are equal. So, fix indices $i \in \{1, \dots, n_1\}$ and $j \in \{1, \dots, n_4\}$. We must show that the i, j-th entry of $(AB)C$ is equal to the i, j-th entry of $A(BC)$.

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We first compute the i, j-th entry of $(AB)C$.

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We first compute the i, j-th entry of $(AB)C$. The i-th row of the $n_1 \times n_3$ matrix AB is

$$\left[\begin{array}{cccc} \sum_{k=1}^{n_2} a_{i,k} b_{k,1} & \sum_{k=1}^{n_2} a_{i,k} b_{k,2} & \dots & \sum_{k=1}^{n_2} a_{i,k} b_{k,n_3} \end{array}\right].$$

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We first compute the i,j-th entry of $(AB)C$. The i-th row of the $n_1 \times n_3$ matrix AB is $\left[\sum_{k=1}^{n_2} a_{i,k} b_{k,1} \quad \sum_{k=1}^{n_2} a_{i,k} b_{k,2} \quad \ldots \quad \sum_{k=1}^{n_2} a_{i,k} b_{k,n_3} \right]$. The j-th column of the $n_3 \times n_4$ matrix C is $\begin{bmatrix} c_{1,j} \\ c_{2,j} \\ \vdots \\ c_{n_3,j} \end{bmatrix}$.

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We first compute the i,j-th entry of $(AB)C$. The i-th row of the $n_1 \times n_3$ matrix AB is $\begin{bmatrix} \sum_{k=1}^{n_2} a_{i,k}b_{k,1} & \sum_{k=1}^{n_2} a_{i,k}b_{k,2} & \ldots & \sum_{k=1}^{n_2} a_{i,k}b_{k,n_3} \end{bmatrix}$. The j-th column of the $n_3 \times n_4$ matrix C is $\begin{bmatrix} c_{1,j} \\ c_{2,j} \\ \vdots \\ c_{n_3,j} \end{bmatrix}$.

So, the i,j-th entry of the $n_1 \times n_4$ matrix $(AB)C$ is

$$\sum_{\ell=1}^{n_3} \Big((\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell})c_{\ell,j} \Big).$$

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We now compute the i, j-th entry of $A(BC)$.

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We now compute the i,j-th entry of $A(BC)$. The i-th row of the $n_1 \times n_2$ matrix A is $\left[\begin{array}{cccc} a_{i,1} & a_{i,2} & \dots & a_{i,n_2} \end{array}\right]$.

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We now compute the i,j-th entry of $A(BC)$. The i-th row of the $n_1 \times n_2$ matrix A is $\begin{bmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n_2} \end{bmatrix}$. The j-th column of the $n_2 \times n_4$ matrix

BC is $\begin{bmatrix} \sum_{k=1}^{n_3} b_{1,k}c_{k,j} \\ \sum_{k=1}^{n_3} b_{2,k}c_{k,j} \\ \vdots \\ \sum_{k=1}^{n_3} b_{n_2,k}c_{k,j} \end{bmatrix}$.

(g) $(AB)C = A(BC)$

Proof of (g) (continued). We now compute the i, j-th entry of $A(BC)$. The i-th row of the $n_1 \times n_2$ matrix A is $\begin{bmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n_2} \end{bmatrix}$. The j-th column of the $n_2 \times n_4$ matrix

$$BC \text{ is } \begin{bmatrix} \sum\limits_{k=1}^{n_3} b_{1,k} c_{k,j} \\ \sum\limits_{k=1}^{n_3} b_{2,k} c_{k,j} \\ \vdots \\ \sum\limits_{k=1}^{n_3} b_{n_2,k} c_{k,j} \end{bmatrix}.$$

So, the i, j-th entry of the $n_1 \times n_4$ matrix $(AB)C$ is

$$\sum_{\ell=1}^{n_2} \Big(a_{i,\ell} \big(\sum_{k=1}^{n_3} b_{\ell,k} c_{k,j} \big) \Big).$$

(g) $(AB)C = A(BC)$

Proof of (g) (continued). Reminder:

- the i, j-th entry of $(AB)C$ is $\sum_{\ell=1}^{n_3} \Big((\sum_{k=1}^{n_2} a_{i,k} b_{k,\ell}) c_{\ell,j} \Big)$;
- the i, j-th entry of $A(BC)$ is $\sum_{\ell=1}^{n_2} \Big(a_{i,\ell} (\sum_{k=1}^{n_3} b_{\ell,k} c_{k,j}) \Big)$.

(g) $(AB)C = A(BC)$

Proof of (g) (continued). Reminder:

- the i,j-th entry of $(AB)C$ is $\sum_{\ell=1}^{n_3}\Big((\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell})c_{\ell,j}\Big)$;
- the i,j-th entry of $A(BC)$ is $\sum_{\ell=1}^{n_2}\Big(a_{i,\ell}(\sum_{k=1}^{n_3} b_{\ell,k}c_{k,j})\Big)$.

It now remains to show that

$$\sum_{\ell=1}^{n_3}\Big((\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell})c_{\ell,j}\Big) \quad = \quad \sum_{\ell=1}^{n_2}\Big(a_{i,\ell}(\sum_{k=1}^{n_3} b_{\ell,k}c_{k,j})\Big).$$

(g) $(AB)C = A(BC)$

Proof of (g) (continued). Reminder:

- the i,j-th entry of $(AB)C$ is $\sum_{\ell=1}^{n_3}\Big((\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell})c_{\ell,j}\Big)$;
- the i,j-th entry of $A(BC)$ is $\sum_{\ell=1}^{n_2}\Big(a_{i,\ell}(\sum_{k=1}^{n_3} b_{\ell,k}c_{k,j})\Big)$.

It now remains to show that

$$\sum_{\ell=1}^{n_3}\Big((\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell})c_{\ell,j}\Big) \quad = \quad \sum_{\ell=1}^{n_2}\Big(a_{i,\ell}(\sum_{k=1}^{n_3} b_{\ell,k}c_{k,j})\Big).$$

For this, we compute (next slide):

(g) $(AB)C = A(BC)$

Proof of (g) (continued).

$$\begin{aligned}
\sum_{\ell=1}^{n_3}\Big(\big(\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell}\big)c_{\ell,j}\Big) &= \sum_{\ell=1}^{n_3}\Big(\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell}c_{\ell,j}\Big) \\
&= \sum_{k=1}^{n_2}\Big(\sum_{\ell=1}^{n_3} a_{i,k}b_{k,\ell}c_{\ell,j}\Big) \\
&= \sum_{k=1}^{n_2}\Big(a_{i,k}\big(\sum_{\ell=1}^{n_3} b_{k,\ell}c_{\ell,j}\big)\Big) \\
&= \sum_{\ell=1}^{n_2}\Big(a_{i,\ell}\big(\sum_{k=1}^{n_3} b_{\ell,k}c_{k,j}\big)\Big)
\end{aligned}$$

and we obtain the equality that we needed. This proves (g). □

Theorem 1.7.5

For any matrices A, B, and C, and any scalars α and β, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalars all belong to the same field $\mathbb{F}$):

(a) $(\alpha + \beta)A = \alpha A + \beta A$;
(b) $(\alpha\beta)A = \alpha(\beta A)$
(c) $A + B = B + A$;
(d) $(A + B) + C = A + (B + C)$;
(e) $(A + B)C = AC + BC$;
(f) $A(B + C) = AB + AC$;
(g) $(AB)C = A(BC)$;
(h) $(\alpha A)B = \alpha(AB)$;
(i) $A(\alpha B) = \alpha(AB)$.

- **Warning:** Matrix multiplication is **not** commutative, that is, for matrices A and B,

$$AB \neq BA.$$

- **Warning:** Matrix multiplication is **not** commutative, that is, for matrices A and B,

$$AB \cancel{=} BA.$$

- In fact, it is possible that one of AB and BA is defined, while the other one is not.
  - For instance, if $A \in \mathbb{F}^{2\times 3}$ and $B \in \mathbb{F}^{3\times 4}$, where $\mathbb{F}$ is some field, then AB is defined, but BA is not.

- **Warning:** Matrix multiplication is **not** commutative, that is, for matrices A and B,

$$AB \not\asymp BA.$$

- In fact, it is possible that one of AB and BA is defined, while the other one is not.
    - For instance, if $A \in \mathbb{F}^{2\times 3}$ and $B \in \mathbb{F}^{3\times 4}$, where $\mathbb{F}$ is some field, then AB is defined, but BA is not.
- Moreover, it is possible that both AB and BA are defined, but are not of the same size.
    - For instance, if $A \in \mathbb{F}^{2\times 3}$ and $B \in \mathbb{F}^{3\times 2}$, where $\mathbb{F}$ is some field, then $AB \in \mathbb{F}^{2\times 2}$ and $BA \in \mathbb{F}^{3\times 3}$.

- **Warning:** Matrix multiplication is **not** commutative, that is, for matrices A and B,

$$AB \xcancel{=} BA.$$

- In fact, it is possible that one of AB and BA is defined, while the other one is not.
  - For instance, if $A \in \mathbb{F}^{2\times 3}$ and $B \in \mathbb{F}^{3\times 4}$, where $\mathbb{F}$ is some field, then AB is defined, but BA is not.
- Moreover, it is possible that both AB and BA are defined, but are not of the same size.
  - For instance, if $A \in \mathbb{F}^{2\times 3}$ and $B \in \mathbb{F}^{3\times 2}$, where $\mathbb{F}$ is some field, then $AB \in \mathbb{F}^{2\times 2}$ and $BA \in \mathbb{F}^{3\times 3}$.
- Finally, it is possible that AB and BA are both defined, and are of the same size, but $AB \neq BA$.
  - For example, for $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, we have that $AB = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ and $BA = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, and so $AB \neq BA$.

Corollary 1.7.6

For any matrices A, B, vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, and scalars α and β, the following hold (provided the matrices and vectors are of compatible size for the operation in question, and the entries of our matrices, the entries of our vectors, and our scalars all belong to the same field $\mathbb{F}$):

(a) $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$;

(b) $(\alpha\beta)\mathbf{u} = \alpha(\beta\mathbf{u})$

(c) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;

(d) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;

(e) $(A + B)\mathbf{u} = A\mathbf{u} + B\mathbf{u}$;

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

(g) $(AB)\mathbf{u} = A(B\mathbf{u})$;

(h) $(\alpha A)\mathbf{u} = \alpha(A\mathbf{u})$;

(i) $A(\alpha\mathbf{u}) = \alpha(A\mathbf{u})$.

- We can define powers of **square** matrices in a natural way, as follows.

- We can define powers of **square** matrices in a natural way, as follows.
- For a field $\mathbb{F}$ and a square matrix $A \in \mathbb{F}^{n \times n}$, we define
  - $A^0 := I_n$;
  - $A^{m+1} := A^m A$ for all non-negative integers m.

- We can define powers of **square** matrices in a natural way, as follows.
- For a field $\mathbb{F}$ and a square matrix $A \in \mathbb{F}^{n \times n}$, we define
  - $A^0 := I_n$;
  - $A^{m+1} := A^m A$ for all non-negative integers m.
- So, by convention, we set $A^0 := I_n$, and for any positive integer m, we have that

$$A^m = \underbrace{A \dots A}_{m},$$

where we did not have to indicate parentheses since, by Theorem 1.7.5, matrix multiplication is associative.

6. The transpose of a matrix

6. The transpose of a matrix

- Given a matrix $A \in \mathbb{F}^{n \times m}$ (where $\mathbb{F}$ is a field), the *transpose* of A, denoted by A^T, is the matrix in $\mathbb{F}^{m \times n}$ s.t. the i, j-th entry of A^T is the j, i-th entry of A, for all indices $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.

6. The transpose of a matrix

- Given a matrix $A \in \mathbb{F}^{n \times m}$ (where $\mathbb{F}$ is a field), the *transpose* of A, denoted by A^T, is the matrix in $\mathbb{F}^{m \times n}$ s.t. the i, j-th entry of A^T is the j, i-th entry of A, for all indices $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
- In other words, to form A^T from A, the columns of A (from left to right) become the rows of A^T (from top to bottom), and likewise, the rows of A (from top to bottom) become the columns of A^T (from left to right).

$$A = \begin{bmatrix} \blacklozenge & * & * & * & * & \lozenge \\ \blacklozenge & * & * & * & * & \lozenge \\ \blacklozenge & * & * & * & * & \lozenge \\ \blacklozenge & * & * & * & * & \lozenge \end{bmatrix} \longrightarrow A^T = \begin{bmatrix} \blacklozenge & \blacklozenge & \blacklozenge & \blacklozenge \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ \lozenge & \lozenge & \lozenge & \lozenge \end{bmatrix}$$

$$A = \begin{bmatrix} \blacklozenge & * & * & * & * & \lozenge \\ \blacklozenge & * & * & * & * & \lozenge \\ \blacklozenge & * & * & * & * & \lozenge \\ \blacklozenge & * & * & * & * & \lozenge \end{bmatrix} \longrightarrow A^T = \begin{bmatrix} \blacklozenge & \blacklozenge & \blacklozenge & \blacklozenge \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ \lozenge & \lozenge & \lozenge & \lozenge \end{bmatrix}$$

- For example, if $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

- In order to save space, we often specify column vectors in terms of transposes of row vectors.
- For instance, we often write something like

$$\mathbf{u} = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}^T \text{ instead of } \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}.$$

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

a. $(A^T)^T = A$;
b. $(A+B)^T = A^T + B^T$;
c. $(\alpha A)^T = \alpha A^T$
d. $(AB)^T = B^T A^T$.

Proof.

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

a. $(A^T)^T = A$;
b. $(A+B)^T = A^T + B^T$;
c. $(\alpha A)^T = \alpha A^T$
d. $(AB)^T = B^T A^T$.

Proof. Parts (a), (b), and (c) are obvious. Let us prove (d).

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

- (a) $(A^T)^T = A$;
- (b) $(A+B)^T = A^T + B^T$;
- (c) $(\alpha A)^T = \alpha A^T$
- (d) $(AB)^T = B^T A^T$.

Proof. Parts (a), (b), and (c) are obvious. Let us prove (d). Fix matrices $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, and set $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times m}$ and $B = \left[\begin{array}{c} b_{i,j} \end{array}\right]_{m\times p}$.

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

(a) $(A^T)^T = A$;

(b) $(A+B)^T = A^T + B^T$;

(c) $(\alpha A)^T = \alpha A^T$

(d) $(AB)^T = B^T A^T$.

Proof. Parts (a), (b), and (c) are obvious. Let us prove (d). Fix matrices $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, and set $A = \left[\ a_{i,j}\ \right]_{n\times m}$ and $B = \left[\ b_{i,j}\ \right]_{m\times p}$. Clearly, $AB \in \mathbb{F}^{n\times p}$, and so $(AB)^T \in \mathbb{F}^{p\times n}$.

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

(a) $(A^T)^T = A$;

(b) $(A+B)^T = A^T + B^T$;

(c) $(\alpha A)^T = \alpha A^T$

(d) $(AB)^T = B^T A^T$.

Proof. Parts (a), (b), and (c) are obvious. Let us prove (d). Fix matrices $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, and set $A = \left[\begin{array}{c} a_{i,j} \end{array} \right]_{n\times m}$ and $B = \left[\begin{array}{c} b_{i,j} \end{array} \right]_{m\times p}$. Clearly, $AB \in \mathbb{F}^{n\times p}$, and so $(AB)^T \in \mathbb{F}^{p\times n}$. On the other hand, we have that $B^T \in \mathbb{F}^{p\times m}$ and $A^T \in \mathbb{F}^{m\times n}$, and so $B^T A^T \in \mathbb{F}^{p\times n}$.

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

(a) $(A^T)^T = A$;

(b) $(A + B)^T = A^T + B^T$;

(c) $(\alpha A)^T = \alpha A^T$

(d) $(AB)^T = B^T A^T$.

Proof. Parts (a), (b), and (c) are obvious. Let us prove (d). Fix matrices $A \in \mathbb{F}^{n \times m}$ and $B \in \mathbb{F}^{m \times p}$, and set $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times m}$ and $B = \begin{bmatrix} b_{i,j} \end{bmatrix}_{m \times p}$. Clearly, $AB \in \mathbb{F}^{n \times p}$, and so $(AB)^T \in \mathbb{F}^{p \times n}$. On the other hand, we have that $B^T \in \mathbb{F}^{p \times m}$ and $A^T \in \mathbb{F}^{m \times n}$, and so $B^T A^T \in \mathbb{F}^{p \times n}$. So, both $(AB)^T$ and $B^T A^T$ are $p \times n$ matrices with entries in $\mathbb{F}$.

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

a. $(A^T)^T = A$;
b. $(A+B)^T = A^T + B^T$;
c. $(\alpha A)^T = \alpha A^T$
d. $(AB)^T = B^T A^T$.

Proof. Parts (a), (b), and (c) are obvious. Let us prove (d). Fix matrices $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, and set $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times m}$ and $B = \begin{bmatrix} b_{i,j} \end{bmatrix}_{m\times p}$. Clearly, $AB \in \mathbb{F}^{n\times p}$, and so $(AB)^T \in \mathbb{F}^{p\times n}$. On the other hand, we have that $B^T \in \mathbb{F}^{p\times m}$ and $A^T \in \mathbb{F}^{m\times n}$, and so $B^T A^T \in \mathbb{F}^{p\times n}$. So, both $(AB)^T$ and $B^T A^T$ are $p\times n$ matrices with entries in $\mathbb{F}$. It remains to show that the corresponding entries of $(AB)^T$ and $B^T A^T$ are the same.

(d) $(AB)^T = B^T A^T$

Proof of (d) (continued).

(d) $(AB)^T = B^T A^T$

Proof of (d) (continued). Fix indices $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, n\}$; we will show that the i, j-th entry of $(AB)^T$ is equal to the i, j-th entry of $B^T A^T$.

By the definition of matrix transpose, the i, j-th entry of $(AB)^T$ is equal to the j, i-th entry of AB, which is equal to $\sum_{k=1}^{m} a_{j,k} b_{k,i}$.

(d) $(AB)^T = B^T A^T$

Proof of (d) (continued). Fix indices $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, n\}$; we will show that the i, j-th entry of $(AB)^T$ is equal to the i, j-th entry of $B^T A^T$.

By the definition of matrix transpose, the i, j-th entry of $(AB)^T$ is equal to the j, i-th entry of AB, which is equal to $\sum_{k=1}^{m} a_{j,k} b_{k,i}$.

We now compute the i, j-th entry of $B^T A^T$.

(d) $(AB)^T = B^T A^T$

Proof of (d) (continued). Fix indices $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, n\}$; we will show that the i, j-th entry of $(AB)^T$ is equal to the i, j-th entry of $B^T A^T$.

By the definition of matrix transpose, the i, j-th entry of $(AB)^T$ is equal to the j, i-th entry of AB, which is equal to $\sum_{k=1}^{m} a_{j,k} b_{k,i}$.

We now compute the i, j-th entry of $B^T A^T$. We observe that i-th row of the matrix B^T is $\begin{bmatrix} b_{1,i} & b_{2,i} & \dots & b_{m,i} \end{bmatrix}$, whereas the j-th column of the matrix A^T is $\begin{bmatrix} a_{j,1} & a_{j,2} & \dots & a_{j,m} \end{bmatrix}^T$.

(d) $(AB)^T = B^T A^T$

Proof of (d) (continued). Fix indices $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, n\}$; we will show that the i, j-th entry of $(AB)^T$ is equal to the i, j-th entry of $B^T A^T$.

By the definition of matrix transpose, the i, j-th entry of $(AB)^T$ is equal to the j, i-th entry of AB, which is equal to $\sum_{k=1}^{m} a_{j,k} b_{k,i}$.

We now compute the i, j-th entry of $B^T A^T$. We observe that i-th row of the matrix B^T is $\begin{bmatrix} b_{1,i} & b_{2,i} & \dots & b_{m,i} \end{bmatrix}$, whereas the j-th column of the matrix A^T is $\begin{bmatrix} a_{j,1} & a_{j,2} & \dots & a_{j,m} \end{bmatrix}^T$. So, the i, j-th entry of the matrix $B^T A^T$ is

$$b_{1,i} a_{j,1} + b_{2,i} a_{j,2} + \cdots + b_{m,i} a_{j,m} = \sum_{k=1}^{m} b_{k,i} a_{j,k} = \sum_{k=1}^{m} a_{j,k} b_{k,i}.$$

(d) $(AB)^T = B^T A^T$

Proof of (d) (continued). Fix indices $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, n\}$; we will show that the i,j-th entry of $(AB)^T$ is equal to the i,j-th entry of $B^T A^T$.

By the definition of matrix transpose, the i,j-th entry of $(AB)^T$ is equal to the j,i-th entry of AB, which is equal to $\sum_{k=1}^{m} a_{j,k} b_{k,i}$.

We now compute the i,j-th entry of $B^T A^T$. We observe that i-th row of the matrix B^T is $\begin{bmatrix} b_{1,i} & b_{2,i} & \dots & b_{m,i} \end{bmatrix}$, whereas the j-th column of the matrix A^T is $\begin{bmatrix} a_{j,1} & a_{j,2} & \dots & a_{j,m} \end{bmatrix}^T$. So, the i,j-th entry of the matrix $B^T A^T$ is

$$b_{1,i} a_{j,1} + b_{2,i} a_{j,2} + \cdots + b_{m,i} a_{j,m} = \sum_{k=1}^{m} b_{k,i} a_{j,k} = \sum_{k=1}^{m} a_{j,k} b_{k,i}.$$

We have now shown that the corresponding entries of the $p \times n$ matrices $(AB)^T$ and $B^T A^T$ are the same, and we deduce that $(AB)^T = B^T A^T$. This proves (d). $\square$

Proposition 1.8.1

For any matrices A and B, and any scalar α, the following hold (provided the matrices are of compatible size for the operation in question, and the entries of our matrices and our scalar belong to the same field $\mathbb{F}$):

a. $(A^T)^T = A$;
b. $(A+B)^T = A^T + B^T$;
c. $(\alpha A)^T = \alpha A^T$
d. $(AB)^T = B^T A^T$.