Linear Algebra 1: Lecture 3

Irena Penev

Winter 2022/2023

In what follows, $\mathbb{F}$ is a fixed field. For now, we will assume that $\mathbb{F}$ is one of the following: $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$ (for some prime number p). However, all the results we prove in this lecture about matrices with entries in $\mathbb{F}$ will still hold for general fields $\mathbb{F}$ (not just $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}_p$ with prime p), though we cannot fully justify this before properly defining fields.

In the remainder of this lecture, we will use the fact that for all $a, b, c \in \mathbb{F}$, the following are satisfied:

- $a + b = b + a$ (commutativity of addition);
- $(a + b) + c = a + (b + c)$ (associativity of addition);
- $ab = ba$ (commutativity of multiplication);
- $(ab)c = a(bc)$ (associativity of multiplication);
- $a(b+c) = ab+ac$ (distributive property of multiplication over addition);
- $(a+b)c = ac+bc$ (distributive property of multiplication over addition);

The above properties are true for $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}_p$ (where p is a prime number). They are in fact true for all fields $\mathbb{F}$, though we have not seen this formally yet.

1 Matrix addition and multiplication

1.1 Matrix addition and scalar multiplication

Given matrices $A = [a_{i,j}]_{n\times m}$ and $B = [b_{i,j}]_{n\times m}$ in $\mathbb{F}^{n\times m}$, we define

$$A + B \quad := \quad [a_{i,j} + b_{i,j}]_{n\times m}.$$

Thus, we add matrices by adding corresponding entries, i.e.

$$[a_{i,j}]_{n\times m} + [b_{i,j}]_{n\times m} \quad = \quad [a_{i,j} + b_{i,j}]_{n\times m}.$$

Further, given a matrix $A = [a_{i,j}]_{n\times m}$ in $\mathbb{F}^{n\times m}$ and a scalar $c \in \mathbb{F}$, we define

$$cA \;\; := \;\; [ca_{i,j}]_{n\times m}.$$

Thus, cA is the matrix obtained from A by multiplying each entry of A by the scalar c, i.e.

$$c\,[a_{i,j}]_{n\times m} \;\; = \;\; [ca_{i,j}]_{n\times m}\,.$$

.

1.2 Matrix multiplication

Suppose we are given two matrices, $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, where $B = \left[\begin{array}{ccc} \mathbf{b}_1 & \ldots & \mathbf{b}_p \end{array}\right]$.[1] We define

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

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:

$$(n\times m) \quad \cdot \quad (m\times p) \quad = \quad (n\times p).$$

Example 1.1. *Let*

$$A \;\; = \;\; \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \end{bmatrix} \qquad \textit{and} \qquad B \;\; = \;\; \begin{bmatrix} 1 & 1 \\ -2 & 1 \\ 0 & -1 \end{bmatrix},$$

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

Solution. We set

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

so that $B = \left[\begin{array}{cc} \mathbf{b}_1 & \mathbf{b}_2 \end{array}\right]$. Then $AB = \left[\begin{array}{cc} A\mathbf{b}_1 & A\mathbf{b}_2 \end{array}\right]$.

---

[1] So, $\mathbf{b}_1, \ldots, \mathbf{b}_p$ are the columns of B, appearing in that order (from left to right) in B.

We compute

$$\begin{aligned}
A\mathbf{b}_1 &= \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \\
&= 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-2) \begin{bmatrix} 2 \\ -3 \end{bmatrix} + 0 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -4 \\ 6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} -3 \\ 6 \end{bmatrix};
\end{aligned}$$

$$\begin{aligned}
A\mathbf{b}_2 &= \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\
&= 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ -3 \end{bmatrix} + (-1) \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ -3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 4 \\ -4 \end{bmatrix}.
\end{aligned}$$

So,

$$AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ 6 & -4 \end{bmatrix}.$$

□

1.3 Another way to compute the product of two matrices

Suppose we are given a matrix $A \in \mathbb{F}^{n\times m}$ and a vector $\mathbf{v} \in \mathbb{F}^m$. Set $A = [a_{i,j}]_{n\times m}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}$. Then by the definition of a matrix-vector product, we have that $A\mathbf{v} \in \mathbb{F}^n$ and

$$
\begin{aligned}
A\mathbf{v} &= \begin{bmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,k} & \cdots & a_{i,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,k} & \cdots & a_{n,m} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_k \\ \vdots \\ v_m \end{bmatrix} \\
&= v_1 \begin{bmatrix} a_{1,1} \\ \vdots \\ a_{i,1} \\ \vdots \\ a_{n,1} \end{bmatrix} + \cdots + v_k \begin{bmatrix} a_{1,k} \\ \vdots \\ a_{i,k} \\ \vdots \\ a_{n,k} \end{bmatrix} + \cdots + v_m \begin{bmatrix} a_{1,m} \\ \vdots \\ a_{i,m} \\ \vdots \\ a_{n,m} \end{bmatrix} \\
&= \begin{bmatrix} a_{1,1}v_1 + \cdots + a_{1,k}v_k + \cdots + a_{1,m}v_m \\ \vdots \\ a_{i,1}v_1 + \cdots + a_{i,k}v_k + \cdots + a_{i,m}v_m \\ \vdots \\ a_{n,1}v_1 + \cdots + a_{n,k}v_k + \cdots + a_{n,m}v_m \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^{m} a_{1,k}v_k \\ \vdots \\ \sum_{k=1}^{m} a_{i,k}v_k \\ \vdots \\ \sum_{k=1}^{m} a_{n,k}v_k \end{bmatrix}.
\end{aligned}
$$

So, the i-th entry of the vector $A\mathbf{v}$ is $\sum\limits_{k=1}^{m} a_{i,k}v_k$.

Let us now consider the product of two matrices. Suppose we are given matrices $A \in \mathbb{F}^{n\times m}$ and $B \in \mathbb{F}^{m\times p}$, and set $A = [a_{i,j}]_{n\times m}$ and $B = [b_{i,j}]_{m\times p}$. The matrix AB belongs to $\mathbb{F}^{n\times p}$. We would like to determine the i, j-th entry of the matrix AB. This is precisely the i-th entry of the j-th column of AB, and by the definition of matrix product, the j-th entry of AB is the vector $A\mathbf{b}_j$, where $\mathbf{b}_j = \begin{bmatrix} b_{1,j} \\ \vdots \\ b_{m,j} \end{bmatrix}$ is the j-th column of B. Using the formula for the matrix-vector product that we obtained above, we see that the i-th entry of the vector $A\mathbf{b}_j$ is $\sum\limits_{k=1}^{m} a_{i,k}b_{k,j}$. So, the i, j-th entry of the $n \times p$ matrix AB is $\sum\limits_{k=1}^{m} a_{i,k}b_{k,j}$.

Here is a way to visualize the product of two matrices. 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. Schematically, this is represented below. The matrix A is on the bottom-left, the matrix B is on the top-right, and the matrix AB is on the bottom-right (squeezed between A and B). The i-th row of A is in red, the j-th column of B is in blue, and the i, j-th entry of AB is $\sum\limits_{k=1}^{m} a_{i,k}b_{k,j}$.

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

Another way to write this is as follows:

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

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.

Example 1.2. *Let*

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

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).

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

By performing arithmetic (in $\mathbb{Z}_2$) on the entries of the matrix that we obtained, we get:

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

□

Example 1.3. *Let*

$$A \;=\; \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 2 & 4 \end{bmatrix} \qquad \text{and} \qquad B \;=\; \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \end{bmatrix},$$

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

Solution. We compute as shown below.

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} (1 \cdot 1 + 2 \cdot 2) & (1 \cdot 0 + 2 \cdot 3) & (1 \cdot (-1) + 2 \cdot 0) \\ ((-1) \cdot 1 + 0 \cdot 2) & ((-1) \cdot 0 + 0 \cdot 3) & ((-1) \cdot (-1) + 0 \cdot 0) \\ (2 \cdot 1 + 4 \cdot 2) & (2 \cdot 0 + 4 \cdot 3) & (2 \cdot (-1) + 4 \cdot 0) \end{bmatrix}$$

After performing arithmetic on the matrix that we obtained, we get

$$AB \;=\; \begin{bmatrix} 5 & 6 & -1 \\ -1 & 0 & 1 \\ 10 & 12 & -2 \end{bmatrix}.$$

□

2 Properties of matrix addition and multiplication

Theorem 2.1. *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) $A(\alpha B) = \alpha(AB)$.

Proof. Parts (a)-(d) readily follow from the definition of the relevant operations and from the properties of addition and multiplication in the field $\mathbb{F}$ (listed at the beginning of the lecture). We prove (a) to illustrate the principle; the proof of (b)-(d) is left as an exercise. Fix a matrix $A = [a_{i,j}]_{n\times m}$ in $\mathbb{F}^{n\times m}$, and fix scalars $\alpha, \beta \in \mathbb{F}$. Then

$$\begin{array}{rcll}
(\alpha+\beta)A & = & (\alpha+\beta)\,[a_{i,j}]_{n\times m} & \\
& = & [(\alpha+\beta)a_{i,j}]_{n\times m} & \text{by the definition of} \\
& & & \text{scalar-matrix product} \\
& = & [\alpha a_{i,j} + \beta a_{i,j}]_{n\times m} & \text{by the distributive property} \\
& & & \text{of mult. over add. in } \mathbb{F} \\
& = & [\alpha a_{i,j}]_{n\times m} + [\beta a_{i,j}]_{n\times m} & \text{by the definition of} \\
& & & \text{matrix addition} \\
& = & \alpha\,[a_{i,j}]_{n\times m} + \beta\,[a_{i,j}]_{n\times m} & \text{by the definition of} \\
& & & \text{scalar-matrix product} \\
& = & \alpha A + \beta A. &
\end{array}$$

Thus, (a) holds.

Next, we prove (e). Fix matrices $A = [a_{i,j}]_{n\times m}$ and $B = [b_{i,j}]_{n\times m}$ in $\mathbb{F}^{n\times m}$, and fix a matrix $C = [c_{i,j}]_{m\times p}$ in $\mathbb{F}^{m\times p}$. We compute:

$$\begin{array}{rcl}
(A+B)C & = & \Big(\left[\ a_{i,j}\ \right]_{n\times m} + \left[\ b_{i,j}\ \right]_{n\times m}\Big)\left[\ c_{i,j}\ \right]_{m\times p} \\
& = & \left[\ a_{i,j} + b_{i,j}\ \right]_{n\times m}\ \left[\ c_{i,j}\ \right]_{m\times p} \\
& = & \left[\ \sum\limits_{k=1}^{m} (a_{i,k} + b_{i,k})c_{k,j}\ \right]_{n\times p} \\
& \overset{(*)}{=} & \left[\ \Big(\sum\limits_{k=1}^{m} a_{i,k}c_{k,j}\Big) + \Big(\sum\limits_{k=1}^{m} b_{i,k}c_{k,j}\Big)\ \right]_{n\times p} \\
& = & \left[\ \sum\limits_{k=1}^{m} a_{i,k}c_{k,j}\ \right]_{n\times p} + \left[\ \sum\limits_{k=1}^{m} b_{i,k}c_{k,j}\ \right]_{n\times p} \\
& = & AC + BC
\end{array}$$

where (*) follows from the fact that addition distributes over multiplication in the field $\mathbb{F}$, and the rest follows from the appropriate definitions. This proves (e). The proof of (f) is similar.

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

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 $\left[\begin{array}{cccc} \sum\limits_{k=1}^{n_2} a_{i,k}b_{k,1} & \sum\limits_{k=1}^{n_2} a_{i,k}b_{k,2} & \cdots & \sum\limits_{k=1}^{n_2} a_{i,k}b_{k,n_3} \end{array}\right]$. The j-th column of the $n_3 \times n_4$ matrix C is $\left[\begin{array}{c} c_{1,j} \\ c_{2,j} \\ \vdots \\ c_{n_3,j} \end{array}\right]$. So, the i,j-th entry of the $n_1 \times n_4$ matrix $(AB)C$ is $\sum\limits_{\ell=1}^{n_3} \Big((\sum\limits_{k=1}^{n_2} a_{i,k}b_{k,\ell}) c_{\ell,j} \Big)$.

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 $\left[\begin{array}{cccc} a_{i,1} & a_{i,2} & \dots & a_{i,n_2} \end{array}\right]$. The j-th column of the $n_2 \times n_4$ matrix BC is $\left[\begin{array}{c} \sum\limits_{k=1}^{n_3} b_{1,k}c_{k,j} \\ \sum\limits_{k=1}^{n_3} b_{2,k}c_{k,j} \\ \vdots \\ \sum\limits_{k=1}^{n_3} b_{n_2,k}c_{k,j} \end{array}\right]$. So, the i,j-th entry of the $n_1 \times n_4$ matrix $(AB)C$ is $\sum\limits_{\ell=1}^{n_2} \Big(a_{i,\ell} (\sum\limits_{k=1}^{n_3} b_{\ell,k}c_{k,j}) \Big)$.

It now remains to show that $\sum\limits_{\ell=1}^{n_3} \Big((\sum\limits_{k=1}^{n_2} a_{i,k}b_{k,\ell}) c_{\ell,j} \Big) = \sum\limits_{\ell=1}^{n_2} \Big(a_{i,\ell} (\sum\limits_{k=1}^{n_3} b_{\ell,k}c_{k,j}) \Big)$. For this, we compute:

$$\begin{aligned}
\sum_{\ell=1}^{n_3} \Big((\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell}) c_{\ell,j} \Big) &= \sum_{\ell=1}^{n_3} \Big(\sum_{k=1}^{n_2} a_{i,k}b_{k,\ell}c_{\ell,j} \Big) && \text{by the distributive property of mult. over add. in } \mathbb{F} \\
&= \sum_{k=1}^{n_2} \Big(\sum_{\ell=1}^{n_3} a_{i,k}b_{k,\ell}c_{\ell,j} \Big) && \text{by swapping } \textstyle\sum \text{ signs} \\
&= \sum_{k=1}^{n_2} \Big(a_{i,k} (\sum_{\ell=1}^{n_3} b_{k,\ell}c_{\ell,j}) \Big) && \text{by the distributive property of mult. over add. in } \mathbb{F} \\
&= \sum_{\ell=1}^{n_2} \Big(a_{i,\ell} (\sum_{k=1}^{n_3} b_{\ell,k}c_{k,j}) \Big) && \text{by swapping the names of dummy variables } k \text{ and } \ell,
\end{aligned}$$

and we obtain the equality that we needed. Thus, $(AB)C = A(BC)$. This proves (g).

The proof of (h) and (i) is left as an exercise. $\square$

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 $B \in \mathbb{F}^{3\times 4}$, 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 $B \in \mathbb{F}^{3\times 2}$, 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 $AB \neq BA$. Consider, for example, matrices $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then $AB = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, but $BA = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$.

Recall that a vector is simply a matrix with one column. So, the following corollary is simply a special case of Theorem 2.1.

Corollary 2.2. *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) $A(\alpha\mathbf{u}) = \alpha(A\mathbf{u})$.

2.1 Standard basis vectors. The identity matrix I_n and the zero matrix $O_{n\times m}$

The $n \times m$ *zero matrix* in $\mathbb{F}^{n\times m}$, denoted by $O_{n\times m}$, is the $n \times m$ matrix, all of whose entries are 0.[2] Clearly, we have the following:

- for all matrices $A \in \mathbb{F}^{n\times m}$, $A + O_{n\times m} = O_{n\times m} + A = A$;
- for all matrices $A \in \mathbb{F}^{p\times n}$, $AO_{n\times m} = O_{p\times m}$;

---

[2] The 0 is from our field $\mathbb{F}$.

- for all matrices $A \in \mathbb{F}^{m\times p}$, $O_{n\times m}A = O_{n\times p}$.

For each positive integer n and index $i \in \{1, \dots, n\}$, 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.[3] Thus, we have that

$$\mathbf{e}_1^n = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2^n = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n^n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

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

Note that for any matrix $A \in \mathbb{F}^{n\times m}$ and any index $i \in \{1, \dots, m\}$, $A\mathbf{e}_i^m$ is the i-th column of A. Indeed, if we set $A = \begin{bmatrix} \mathbf{a}_1 & \dots \mathbf{a}_{i-1} & \mathbf{a}_i & \mathbf{a}_{i+1} & \dots & \mathbf{a}_m \end{bmatrix}$, then we have that

$$A\mathbf{e}_i^m \quad = \quad 0\mathbf{a}_1 + \dots + 0\mathbf{a}_{i-1} + 1\mathbf{a}_i + 0\mathbf{a}_{i+1} + \dots + 0\mathbf{a}_m \quad = \quad \mathbf{a}_i.$$

Finally, I_n is the matrix in $\mathbb{F}^{n\times n}$ given by

$$I_n \quad := \quad \begin{bmatrix} \mathbf{e}_1^n & \mathbf{e}_2^n & \dots & \mathbf{e}_n^n \end{bmatrix}.$$

In other words, I_n is the $n \times n$ matrix in $\mathbb{F}^{n\times n}$ that has 1's on the main diagonal, and 0's everywhere else, as shown below.[4]

$$I_n \quad = \quad \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}_{n\times n}$$

Proposition 2.3. *For any vector* $\mathbf{v} \in \mathbb{F}^n$*, we have that* $I_n\mathbf{v} = \mathbf{v}$.

---

[3] So, $\mathbf{e}_i^n$ has n entries, its i-th entry is 1 (this is the 1 from the field $\mathbb{F}$), and all its other entries are 0 (this is the zero from the field $\mathbb{F}$).

[4] As before, the 1's and 0's come from the field $\mathbb{F}$.

Proof. Fix $\mathbf{v} \in \mathbb{F}^n$, and set $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. We then compute

$$\begin{aligned} I_n\mathbf{v} &= \begin{bmatrix} \mathbf{e}_1^n & \mathbf{e}_2^n & \dots & \mathbf{e}_n^n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= v_1\mathbf{e}_1^n + v_2\mathbf{e}_2^n + \cdots + v_n\mathbf{e}_n^n \\ &= \begin{bmatrix} v_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ v_n \end{bmatrix} \\ &= \mathbf{v}, \end{aligned}$$

and we are done. ☐

Proposition 2.4. *For any matrix A, the following hold (provided that A and I_n are of compatible size for the operation in question, and the entries of A and I_n all belong to the same field $\mathbb{F}$):*

(a) $I_nA = A$;

(b) $AI_n = A$.

Proof. We first prove (a). Fix a matrix $A \in \mathbb{F}^{n\times m}$, and set $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$\begin{aligned} I_nA &= I_n\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_m \end{bmatrix} \\ &= \begin{bmatrix} I_n\mathbf{a}_1 & I_n\mathbf{a}_2 & \dots & I_n\mathbf{a}_m \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_m \end{bmatrix} && \text{by Proposition 2.3} \\ &= A. \end{aligned}$$

This proves (a).

It remains to prove (b). Fix a matrix $A \in \mathbb{F}^{p\times n}$, and set $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$.

Recall that for all $i \in \{1, \dots, n\}$, we have that $A\mathbf{e}_i^n = \mathbf{a}_i$. We now compute:

$$\begin{aligned} AI_n &= A\begin{bmatrix} \mathbf{e}_1^n & \mathbf{e}_2^n & \dots & \mathbf{e}_n^n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{e}_1^n & A\mathbf{e}_2^n & \dots & A\mathbf{e}_n^n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \\ &= A. \end{aligned}$$

This proves (b). □

3 The transpose of a matrix

Given a matrix $A \in \mathbb{F}^{n\times m}$, the *transpose* of A, denoted by A^T, is the matrix in $\mathbb{F}^{m\times n}$ such that the i,j-th entry of A^T is the j,i-th entry of A. 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 columns of A^T (from left to right).

Example 3.1. *Let* $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. *Compute* A^T.

Solution.

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

□

Note that vectors are often given in terms of transposes, in order to save space. For instance, we often write something like $\mathbf{u} = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}^T$ instead of $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$.

Proposition 3.2. *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 $A = [a_{i,j}]_{n\times m}$ and $B = [b_{i,j}]_{m\times p}$. 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 $\mathbb{F}$. We need only show that the corresponding entries of $(AB)^T$ and $B^T A^T$ are the same. 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 $\sum\limits_{k=1}^{m} a_{j,k} b_{k,i}$.

We now compute the i, j-th entry of $B^T A^T$. The i-th row of B^T is $\begin{bmatrix} b_{1,i} & b_{2,i} & \dots & b_{m,i} \end{bmatrix}$,[5] and the j-th column of A^T is $\begin{bmatrix} a_{j,1} & a_{j,2} & \dots & a_{j,m} \end{bmatrix}^T$.[6] So, the i, j-th entry of $B^T A^T$ is $b_{1,i} a_{j,1} + b_{2,i} a_{j,2} + \dots + b_{m,i} a_{j,m} = \sum\limits_{k=1}^{m} a_{j,k} b_{k,i}$.

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). □

---

[5] Indeed, this is the transpose of the i-th column of B.

[6] Indeed, this is the transpose of the j-th row of A.