Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-09-23

Darij Grinberg

December 5, 2019

Course website: `https://www.cip.ifi.lmu.de/~grinberg/t/19fla/index.html`

1. Introduction to matrices

This chapter follows [lina, Chapter 2], but we will give a lot fewer details, as many of you will have seen this material already. We will spend half the time introducing notations and half the time stating basic facts.

1.1. Matrices and their entries

This section follows [lina, §2.1].

From now on, $\mathbb{N}$ means the set $\{0, 1, 2, \ldots\}$.

If $n, m \in \mathbb{N}$, then an $n \times m$*-matrix* will mean a rectangular table with n rows and m columns, such that each cell is filled with a number.

"Number" means real number unless stated otherwise.

For example, $\begin{pmatrix} 1 & 7 & 2 \\ -\sqrt{2} & 6 & 1/3 \end{pmatrix}$ is a 2×3-matrix.

A *matrix* just means an $n \times m$-matrix for some n and m.

The *dimensions* of an $n \times m$-matrix are the two integers n and m. We say that a matrix has *size* $n \times m$ if it is an $n \times m$-matrix.

If A is an $n \times m$-matrix, and $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$, then $A_{i,j}$ shall mean the entry of A in row i and column j. This is also called the (i, j)*-th entry* of A.

For example,

$$\begin{pmatrix} 1 & 7 & 2 \\ -\sqrt{2} & 6 & 1/3 \end{pmatrix}_{1,2} = 7;$$
$$\begin{pmatrix} 1 & 7 & 2 \\ -\sqrt{2} & 6 & 1/3 \end{pmatrix}_{2,3} = 1/3.$$

This notation is not quite standard. You will often see people denote the (i,j)-th entry of A by $a_{i,j}$ (using the lowercase version of the letter). I prefer to call it $A_{i,j}$.

1.2. The matrix builder notation

This section follows [lina, §2.2].

Let $n, m \in \mathbb{N}$. Assume that you are given some number $a_{i,j}$ for each pair (i,j) of an $i \in \{1,2,\ldots,n\}$ and a $j \in \{1,2,\ldots,m\}$. Then, $\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$ denotes the $n \times m$-matrix whose (i,j)-th entry is $a_{i,j}$ for all i and j. In other words,

$$\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix}.$$

This is called *matrix builder notation*[1].

Some examples:

$$(i-j)_{1\leq i\leq 2,\ 1\leq j\leq 3} = \begin{pmatrix} 1-1 & 1-2 & 1-3 \\ 2-1 & 2-2 & 2-3 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \end{pmatrix};$$

$$(j-i)_{1\leq i\leq 2,\ 1\leq j\leq 3} = \begin{pmatrix} 1-1 & 2-1 & 3-1 \\ 1-2 & 2-2 & 3-2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix}.$$

The letters i and j in the notation "$\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$" are **dummy variables**, like for example the letter x in $\{x \in \mathbb{R} \mid x > 2\}$. Any two symbols can be used instead:

$$\begin{aligned}(i-j)_{1\leq i\leq 2,\ 1\leq j\leq 3} &= (x-y)_{1\leq x\leq 2,\ 1\leq y\leq 3} = (a-b)_{1\leq a\leq 2,\ 1\leq b\leq 3} \\ &= (j-i)_{1\leq j\leq 2,\ 1\leq i\leq 3}.\end{aligned}$$

The fact that j is before i in the subscript of "$(j-i)_{1\leq j\leq 2,\ 1\leq i\leq 3}$" tells us that j indexes the rows and i the columns, so this is not the same as $(j-i)_{1\leq i\leq 2,\ 1\leq j\leq 3}$.

One simple observation:

Proposition 1.2.1. If A is any $n \times m$-matrix, then

$$\left(A_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m} = A.$$

1.3. Row and column vectors

This section follows [lina, §2.3].

[1] in analogy to the *set builder notation* $\{b_i \mid i \in \{1,2,\ldots,n\}\} = \{b_1, b_2, \ldots, b_n\}$ for the set consisting of n given objects $b_1, b_2, \ldots, b_n$

Definition 1.3.1. Let $n \in \mathbb{N}$.
A **row vector of size** n means a $1 \times n$-matrix.
A **column vector of size** n means an $n \times 1$-matrix.

For example, $\begin{pmatrix} a & b \end{pmatrix}$ is a row vector of size 2, while $\begin{pmatrix} a \\ b \end{pmatrix}$ is a column vector of size 2.
The *j-th entry* of a row vector A is $A_{1,j}$.
The *j-th entry* of a column vector A is $A_{j,1}$.

Definition 1.3.2. Let $n \in \mathbb{N}$. We let $\mathbb{R}^n$ denote the set of all column vectors of size n (with real entries).

1.4. Transposes

This section follows [lina, §2.4].

Definition 1.4.1. The **transpose** of an $n \times m$-matrix A is defined to be the $m \times n$-matrix $\left(A_{j,i}\right)_{1 \leq i \leq m,\ 1 \leq j \leq n}$. It is denoted by A^T.

For example,

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}^T = \begin{pmatrix} a & a' \\ b & b' \\ c & c' \end{pmatrix};$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}^T = \begin{pmatrix} a & b & c \end{pmatrix};$$

$$\begin{pmatrix} a & b & c \end{pmatrix}^T = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

This allows us to use transposes as space-saving devices: Instead of writing $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, you can just write $\begin{pmatrix} a & b & c \end{pmatrix}^T$.

Proposition 1.4.2. Let $n, m \in \mathbb{N}$. Let A be an $n \times m$-matrix. Then, $\left(A^T\right)^T = A$.

1.5. Addition, scaling and multiplication

This section follows [lina, §2.5].

Definition 1.5.1. Let A and B be two matrices of the same dimension (i.e., same number of rows & same number of columns). Then, $A+B$ denotes the matrix obtained by adding A and B entry by entry (i.e., adding each entry of A to the corresponding entry of B).

In formulas: If A and B are two $n \times m$-matrices, then

$$A + B = \left(A_{i,j} + B_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}.$$

For example,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix}.$$

Definition 1.5.2. Let A be a matrix. Let λ be a number. Then, λA denotes the matrix obtained by multiplying each entry of A by λ.

In formulas: If A is an $n \times m$-matrix and λ is a number, then

$$\lambda A = \left(\lambda A_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}.$$

For example,

$$\lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}.$$

The operation of computing λA from A is called **scaling** the matrix A by λ.

Definition 1.5.3. Let A and B be two matrices of the same dimensions. Then, $A - B$ denotes the matrix $A + (-1)B$.

For example,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a-a' & b-b' \\ c-c' & d-d' \end{pmatrix}.$$

Definition 1.5.4. Let $n, m, p \in \mathbb{N}$. Let A be an $n \times m$-matrix. Let B be an $m \times p$-matrix. Then, the product AB of these two matrices is defined as follows:

$$AB = \left(A_{i,1}B_{1,j} + A_{i,2}B_{2,j} + \cdots + A_{i,m}B_{m,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq p}.$$

This is an $n \times p$-matrix.

Examples:

$$\begin{pmatrix} a & b \\ a' & b' \end{pmatrix} \begin{pmatrix} x & x' \\ y & y' \end{pmatrix} = \begin{pmatrix} ax+by & ax'+by' \\ a'x+b'y & a'x'+b'y' \end{pmatrix};$$

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax+by+cz \\ a'x+b'y+c'z \end{pmatrix};$$

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \end{pmatrix};$$

$$\begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} ax & ay \\ bx & by \end{pmatrix};$$

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \text{undefined}.$$

Examples:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Examples:

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} a+b+c & a+b+c & a+b+c \\ a'+b'+c' & a'+b'+c' & a'+b'+c' \\ a''+b''+c'' & a''+b''+c'' & a''+b''+c'' \end{pmatrix};$$

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a+b+c \\ a'+b'+c' \\ a''+b''+c'' \end{pmatrix};$$

$$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = \begin{pmatrix} a+a'+a'' & b+b'+b'' & c+c'+c'' \end{pmatrix}.$$

1.6. The matrix product rewritten

This section follows [lina, §2.6].

Definition 1.6.1. Let A be an $n \times m$-matrix.

(a) If $i \in \{1, 2, \ldots, n\}$, then $\operatorname{row}_i A$ will mean the i-th row of A. This is a row vector of size m (that is, a $1 \times m$-matrix), and is formally defined as

$$\left(A_{i,y}\right)_{1 \leq x \leq 1,\ 1 \leq y \leq m} = \begin{pmatrix} A_{i,1} & A_{i,2} & \cdots & A_{i,m} \end{pmatrix}.$$

(b) If $j \in \{1,2,\ldots,m\}$, then $\operatorname{col}_j A$ will mean the j-th column of A. This is a column vector of size n (that is an $n \times 1$-matrix), and is formally defined as

$$\left(A_{x,j}\right)_{1\leq x\leq n,\ 1\leq y\leq 1} = \begin{pmatrix} A_{1,j} \\ A_{2,j} \\ \vdots \\ A_{n,j} \end{pmatrix}.$$

For example: If $A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$, then $\operatorname{row}_2 A = \begin{pmatrix} d & e & f \end{pmatrix}$ and $\operatorname{col}_2 A = \begin{pmatrix} b \\ e \end{pmatrix}$.

Notice how the product of two matrices looks like if the first matrix is a row vector and the second is a column vector:

$$\begin{pmatrix} r_1 & r_2 & \cdots & r_m \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} r_1c_1 + r_2c_2 + \cdots + r_mc_m \end{pmatrix}.$$

We shall equate 1×1-matrices with their unique entries, so this becomes

$$\begin{pmatrix} r_1 & r_2 & \cdots & r_m \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = r_1c_1 + r_2c_2 + \cdots + r_mc_m.$$

Now, a collection of formulas for the product of two matrices ([lina, Proposition 2.19]). Note that these formulas are all essentially saying the same thing, but from different points of view, and that's useful.

Proposition 1.6.2. Let $n, m, p \in \mathbb{N}$. Let A be an $n \times m$-matrix, and B be an $m \times p$-matrix.

(a) For every $i \in \{1,2,\ldots,n\}$ and $j \in \{1,2,\ldots,p\}$, we have

$$(AB)_{i,j} = A_{i,1}B_{1,j} + A_{i,2}B_{2,j} + \cdots + A_{i,m}B_{m,j}.$$

(b) For every $i \in \{1,2,\ldots,n\}$ and $j \in \{1,2,\ldots,p\}$, we have

$$(AB)_{i,j} = \operatorname{row}_i A \cdot \operatorname{col}_j B \qquad \left(\text{this means } (\operatorname{row}_i A) \cdot (\operatorname{col}_j B)\right).$$

Thus,

$$AB = \begin{pmatrix} \operatorname{row}_1 A \cdot \operatorname{col}_1 B & \operatorname{row}_1 A \cdot \operatorname{col}_2 B & \cdots & \operatorname{row}_1 A \cdot \operatorname{col}_p B \\ \operatorname{row}_2 A \cdot \operatorname{col}_1 B & \operatorname{row}_2 A \cdot \operatorname{col}_2 B & \cdots & \operatorname{row}_2 A \cdot \operatorname{col}_p B \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{row}_n A \cdot \operatorname{col}_1 B & \operatorname{row}_n A \cdot \operatorname{col}_2 B & \cdots & \operatorname{row}_n A \cdot \operatorname{col}_p B \end{pmatrix}.$$

(c) For every $i \in \{1, 2, \ldots, n\}$, we have

$$\text{row}_i (AB) = (\text{row}_i A) \cdot B.$$

(d) For every $j \in \{1, 2, \ldots, p\}$, we have

$$\text{col}_j (AB) = A \cdot \text{col}_j B.$$

Let us illustrate part **(d)** on an example (with $n = 2$, $m = 2$, $p = 2$, $A = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$ and $B = \begin{pmatrix} x & x' \\ y & y' \end{pmatrix}$):

$$\underbrace{\begin{pmatrix} a & b \\ a' & b' \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x & x' \\ y & y' \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} ax + by & ax' + by' \\ a'x + b'y & a'x' + b'y' \end{pmatrix}}_{AB};$$

$$\underbrace{\begin{pmatrix} ax' + by' \\ a'x' + b'y' \end{pmatrix}}_{\text{col}_j(AB)} = \underbrace{\begin{pmatrix} a & b \\ a' & b' \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} x' \\ y' \end{pmatrix}}_{\text{col}_j B}.$$

1.7. Properties of matrix operations

This section follows [lina, §2.7 and §2.8].

Addition, scaling and multiplication of matrices has the following properties ([lina, Proposition 2.20]):

Proposition 1.7.1. Let $n, m \in \mathbb{N}$.

(a) We have $A + B = B + A$ for any two $n \times m$-matrices A and B.

(b) We have $A + (B + C) = (A + B) + C$ for any three $n \times m$-matrices A, B and C.

(c$_1$) We have $\lambda (A + B) = \lambda A + \lambda B$ for any number λ and any two $n \times m$-matrices A and B.

(c$_2$) We have $(\lambda + \mu) A = \lambda A + \mu A$ for any numbers λ and μ and any $n \times m$-matrix A.

(c$_3$) We have $1A = A$ for any $n \times m$-matrix A.

Furthermore, let $p \in \mathbb{N}$.

(d) We have $A (B + C) = AB + AC$ whenever A is an $n \times m$-matrix and B and C are two $m \times p$-matrices.

(e) We have $(A + B) C = AC + BC$ whenever A and B are two $n \times m$-matrices and C is an $m \times p$-matrix.

(f) We have $\lambda (AB) = (\lambda A) B = A (\lambda B)$ whenever λ is a number and A is an $n \times m$-matrix and B is an $m \times p$-matrix.

Finally, let $q \in \mathbb{N}$.

(g) We have $(AB)C = A(BC)$ whenever A is an $n \times m$-matrix, B is an $m \times p$-matrix and C is a $p \times q$-matrix.

Here is an example for Proposition 1.7.1 **(g)**:

Example 1.7.2. Let $n = 1$ and $m = 3$ and $p = 3$ and $q = 1$, and let

$$A = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \qquad \text{and} \qquad C = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Then,

$$BC = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a+b+c \\ a'+b'+c' \\ a''+b''+c'' \end{pmatrix};$$

$$AB = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = \begin{pmatrix} a+a'+a'' & b+b'+b'' & c+c'+c'' \end{pmatrix}.$$

Now,

$$\begin{aligned} A(BC) &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a+b+c \\ a'+b'+c' \\ a''+b''+c'' \end{pmatrix} \\ &= (a+b+c) + (a'+b'+c') + (a''+b''+c'') \end{aligned}$$

and

$$\begin{aligned} (AB)C &= \begin{pmatrix} a+a'+a'' & b+b'+b'' & c+c'+c'' \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= (a+a'+a'') + (b+b'+b'') + (c+c'+c''), \end{aligned}$$

and these are the same number (namely, the sum of all entries of B).

Proposition 1.7.1 **(g)** is called *associativity of matrix multiplication* and shows that a product ABC of three matrices is unambiguous (i.e., the result does not depend on whether we interpret it as $(AB)C$ or as $A(BC)$).

The same holds for products of four matrices:

$$((AB)C)D = (A(BC))D = A((BC)D) = (AB)(CD) = A(B(CD)).$$

So we can write $ABCD$ without worrying about ambiguity.

More generally:

Proposition 1.7.3. Let $A_1, A_2, \ldots, A_n$ be n matrices such that all the $n-1$ products $A_i A_{i+1}$ are well-defined (i.e., the number of columns of A_i equals the number of rows of A_{i+1}).

Then, the product $A_1 A_2 \cdots A_n$ is unambiguous (i.e., the result does not depend on where you start multiplying it out).

Something similar holds for sums: $A_1 + A_2 + \cdots + A_n$ is unambiguous whenever $A_1, A_2, \ldots, A_n$ are matrices of the same dimensions.

However: Products of matrices **cannot** be reordered at will! (unlike products of numbers). In other words, matrix multiplication is **not commutative**. More precisely, if A and B are two matrices, it can happen that

- AB is defined but BA is not;
- AB and BA are both defined but not equal.

For example, if $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, then $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, so $AB \neq BA$.

References

[lina] Darij Grinberg, *Notes on linear algebra*, version of 13 December 2016. `https://github.com/darijgr/lina`

[Strickland] Neil Strickland, *Linear Algebra for Applications - MAS201*, lecture notes, version with edits by myself. `http://www.cip.ifi.lmu.de/~grinberg/t/19fla/MAS201.pdf` See also Neil Strickland's course page `https://neil-strickland.staff.shef.ac.uk/courses/MAS201/` for exercises with solutions.