

Linear Algebra Notes — Lecture 9/8

One-to-One and Onto in Practice, Composition, and Matrix Algebra

Lecture pages 47–59 · Lay §1.9, §2.1

Contents

Overview	1
1 A worked example: a map $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ (pp. 47–48)	2
2 The dictionary: matrices versus linear maps (p. 49)	4
3 Composition and the matrix product (pp. 50–52)	4
4 A formula for the product: column by column (p. 53)	7
5 The row–column rule (pp. 54–55)	8
6 Matrices are not numbers (pp. 55–56)	9
7 Powers of diagonal matrices (pp. 56–57)	10
8 Sums and scalar multiples of matrices (pp. 57–58)	11
9 The algebra rules (p. 59)	12
Summary	13
Corrections to the handwritten notes	14

Overview

Last time (9/3) we learned that a linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the same thing as an $m \times n$ matrix A , via $T(x) = Ax$, and we found pivot tests for “onto” and “one-to-one”. This lecture uses those tests on an example, and then asks what *composing* two maps does to their matrices. The answer defines the matrix product.

1. A worked example: find the standard matrix of a map $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ and test it for one-to-one and onto (p. 47).

2. A “dictionary” that translates between matrices and linear maps (p. 49).
3. Composition $S \circ T$ of linear maps is linear. Its standard matrix is, by definition, the product BA (pp. 50–52).
4. Two ways to compute a product: column by column, $BA = [Ba_1 \cdots Ba_n]$ (p. 53), and entry by entry with the row–column rule (pp. 54–55).
5. Matrix multiplication behaves differently from multiplication of numbers: $AB \neq BA$ in general, $AB = 0$ does not force $A = 0$ or $B = 0$, and cancellation fails (pp. 55–56).
6. Powers of diagonal matrices (pp. 56–57), sums and scalar multiples of matrices (pp. 57–58), and the algebra rules (p. 59).

Every blank and question in the notes is answered in a blue box. Typos are listed at the end.

Notation used throughout.

- $e_1, \dots, e_n$: the standard basis vectors of $\mathbb{R}^n$, the columns of the identity matrix I_n (9/3 notes, p. 40).
- $A = [a_1 \cdots a_n]$: the matrix A written by its columns $a_1, \dots, a_n$.
- a_{ij} : the entry of A in row i , column j ; $A = (a_{ij})$ (p. 54).
- $S \circ T$: the composition “first T , then S ”, $(S \circ T)(x) = S(T(x))$.

1 A worked example: a map $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ (pp. 47–48)

This example runs the whole 9/3 toolkit on one map: it is linear, its standard matrix is 3×2 , and its pivots show it is one-to-one but not onto.

The map is

$$T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix}.$$

It takes vectors with 2 entries to vectors with 3 entries, so the **domain of T is $\mathbb{R}^2$** and the **codomain of T is $\mathbb{R}^3$** .

Is T linear? Yes. Its standard matrix is $A = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}$

Recipe (9/3 notes, p. 41): if T is linear, its standard matrix is $A = [T(e_1) \ T(e_2)]$, the matrix whose j -th column is $T(e_j)$.

Plug in $e_1 = (1, 0)$: $T(e_1) = (3 \cdot 1 + 0, 5 \cdot 1 + 0, 1 + 0) = (3, 5, 1)$. Plug in $e_2 = (0, 1)$: $T(e_2) = (1, 7, 3)$. So

$$A = [T(e_1) \ T(e_2)] = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}.$$

Why T is linear. Multiplying out,

$$\begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 7 \\ 3 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} = T(x),$$

which is the formula we started with (the notes' double arrow on p. 47 links these two expressions). So $T(x) = Ax$ for every x , and every map of the form $x \mapsto Ax$ is linear (9/3 notes, p. 37).

Check: $T(2, -1) = (6 - 1, 10 - 7, 2 - 3) = (5, 3, -1)$ and $A(2, -1)^T = 2(3, 5, 1) - (1, 7, 3) = (5, 3, -1)$. ✓

Row reduction (p. 47). To test one-to-one and onto we need the pivots of A . The notes swap rows 1 and 3 (so the top-left entry is 1), then clear the first column:

$$\begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 3 & 1 \end{bmatrix} \xrightarrow{\substack{R_2 - 5R_1 \\ R_3 - 3R_1}} \begin{bmatrix} 1 & 3 \\ 0 & -8 \\ 0 & -8 \end{bmatrix} \xrightarrow{\substack{-\frac{1}{8}R_2 \\ R_3 - R_2}} \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

(The last step divides row 2 by -8 and then subtracts the new row 2 from row 3; the notes show only the result.) There are pivots in row 1, column 1 and row 2, column 2. Row 3 has no pivot.

Is T one-to-one? Yes. Is T onto? No

One-to-one: yes. Both columns of A contain a pivot. By the theorem on p. 46 (9/3 notes), “each column of A is pivotal” is equivalent to “ T is one-to-one”. Concretely, $Ax = 0$ has only the solution $x = 0$: the reduced system says $x_2 = 0$ and $x_1 + 3x_2 = 0$, so $x_1 = 0$ too.

Onto: no. Row 3 of the echelon form has no pivot. By the theorem on p. 45 (9/3 notes), T is onto exactly when every row of A is pivotal, so T is not onto.

A concrete b that is missed. The image of T is the span of the columns $(3, 5, 1)$ and $(1, 7, 3)$ (9/3 notes, p. 33), a plane in $\mathbb{R}^3$. Every vector (u, v, w) in that plane satisfies $u - v + 2w = 0$ (both columns do: $3 - 5 + 2 = 0$ and $1 - 7 + 6 = 0$). The vector $b = (1, 0, 0)$ gives $1 - 0 + 0 = 1 \neq 0$, so $Ax = b$ has no solution.

Check: row reducing $[A \mid b]$ with $b = (1, 0, 0)$ gives a pivot in the last column (computed numerically), i.e. an equation $0 = \text{nonzero}$: inconsistent. ✓

Why it makes sense: two columns can hold at most two pivots, but onto would need a pivot in each of the three rows. A map $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ can never be onto. (Similarly, a map $\mathbb{R}^3 \rightarrow \mathbb{R}^2$ can never be one-to-one.)

Page 48 is blank in the handwritten notes (only the page number is written). Nothing is missing from this document.

2 The dictionary: matrices versus linear maps (p. 49)

Every statement about an $m \times n$ matrix A has a twin statement about the linear map $T(x) = Ax$, and the professor collects them in one table.

Matrices	Linear maps
$A = [T(e_1) \cdots T(e_n)]$	$T(x) = Ax$
A has n columns	domain of T is $\mathbb{R}^n$
A has m rows	codomain of T is $\mathbb{R}^m$
$Ax = b$	$T(x) = b$
$\text{span}\{T(e_1), \dots, T(e_n)\}$ $\text{span}\{\text{columns of } A\} = \{Ax : x \in \mathbb{R}^n\}$	$= \text{image of } T = \{T(x) : x \in \mathbb{R}^n\}$
A has a pivot in each row	T is onto
A has a pivot in each column	T is one-to-one
?	$S \circ T$

Where each row comes from: the first row is the standard-matrix theorem (9/3 notes, p. 41); rows 2–3 hold because Ax needs x to have one entry per column of A and produces one entry per row; the image row is the theorem “image of $T = \text{span of the columns}$ ” (9/3 notes, p. 33); the pivot rows are the theorems on pp. 45–46 (9/3 notes).

The “?”: composition $S \circ T$ corresponds to the matrix product BA

If $T(x) = Ax$ and $S(y) = By$, then $S \circ T$ is again linear (Section 3) and its standard matrix is the product BA . In fact, p. 51 *defines* BA this way, and p. 53 gives the formula $BA = [Ba_1 \cdots Ba_n]$. So the missing dictionary entry is

$$\text{matrix product } BA \quad \longleftrightarrow \quad \text{composition } S \circ T.$$

Note the order: $S \circ T$ means “do T first”, and BA has A (the matrix of T) on the right, next to x : $(S \circ T)(x) = B(Ax)$.

3 Composition and the matrix product (pp. 50–52)

Doing one linear map after another gives a linear map, and the matrix of that composition is what we call the product of the two matrices.

Theorem (p. 50)

The composition of linear maps is linear: if $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $S : \mathbb{R}^m \rightarrow \mathbb{R}^k$ are linear, then so is $S \circ T : \mathbb{R}^n \rightarrow \mathbb{R}^k$, $(S \circ T)(x) = S(T(x))$.

Proof (p. 50). The setup is

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{T} & \mathbb{R}^m & \xrightarrow{S} & \mathbb{R}^k. \\ & \searrow & \text{ } & \nearrow & \\ & S \circ T & & & \end{array}$$

(The diagram shows the two maps in a row; the curved arrow is “go along both”.) A map f is linear if and only if $f(c_1v_1 + c_2v_2) = c_1f(v_1) + c_2f(v_2)$ for all vectors v_1, v_2 and scalars c_1, c_2 (9/3 notes, p. 38, with $k = 2$). Check this for $f = S \circ T$:

$$S(T(c_1v_1 + c_2v_2)) \stackrel{T \text{ linear}}{=} S(c_1T(v_1) + c_2T(v_2)) \stackrel{S \text{ linear}}{=} c_1S(T(v_1)) + c_2S(T(v_2)).$$

The first step uses linearity of T on v_1, v_2 ; the second uses linearity of S on the vectors $T(v_1), T(v_2)$.

□

Now let A and B be the matrices with $T(x) = Ax$ and $S(y) = By$.

What are the sizes of A and B ? A is $m \times n$, B is $k \times m$

By the dictionary, the number of columns is the dimension of the domain and the number of rows is the dimension of the codomain.

- $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, so A has n columns and m rows: A is $m \times n$.
- $S : \mathbb{R}^m \rightarrow \mathbb{R}^k$, so B has m columns and k rows: B is $k \times m$.
- $S \circ T : \mathbb{R}^n \rightarrow \mathbb{R}^k$, so its standard matrix is $k \times n$. The notes match this with BA , written $(k \times m)(m \times n)$: the inner sizes m agree and the outer sizes $k \times n$ survive.

Definition (product of matrices, p. 51)

The product BA is the standard matrix of $S \circ T$. In other words, BA is the matrix with $(BA)x = B(Ax)$ for all x .

What is the size of BA ? $k \times n$

$S \circ T$ goes from $\mathbb{R}^n$ to $\mathbb{R}^k$, so its standard matrix has n columns and k rows. Check with p. 53's example: a $(2 \times 2)(2 \times 2)$ product is 2×2 . ✓

Does AB also make sense? Yes if $k = n$, no if $k \neq n$

A is $m \times n$ and B is $k \times m$. The product AB means “first S , then T ”: S lands in $\mathbb{R}^k$, but T only accepts inputs from $\mathbb{R}^n$. So $T \circ S$ makes sense exactly when $k = n$. In matrix terms: the number of columns of A (n) must equal the number of rows of B (k). When $k = n$, AB is $(m \times n)(n \times m) = m \times m$.

Example: A is 2×3 and B is 4×2 ($m = 2, n = 3, k = 4$). Then BA is 4×3 , but AB would need $3 = 4$: not defined.

Standard matrix of $T(x) = x$: the identity matrix I_n

Use the recipe $[T(e_1) \cdots T(e_n)]$: here $T(e_j) = e_j$, so the matrix is $[e_1 \cdots e_n] = I_n$, the $n \times n$

matrix with 1's on the diagonal and 0's elsewhere (9/3 notes, p. 34):

$$I_n x = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + \cdots + x_n e_n = x. \checkmark$$

Examples (p. 52): the identity, associativity, commutativity

Why $AI_n = A$ and $I_m A = A$ (for A of size $m \times n$)

$AI_n = A$. Let $S(x) = Ax$ and let $T(x) = x$ be the identity map of $\mathbb{R}^n$, whose matrix is I_n . Then $S \circ T = S$, because $S(T(x)) = S(x)$. Taking standard matrices of both sides gives $AI_n = A$.

$I_m A = A$. Same argument with the identity map of $\mathbb{R}^m$ applied *after* $x \mapsto Ax$. Note the sizes: on the right we need I_n (n = number of columns of A), on the left I_m (m = number of rows). For a square $n \times n$ matrix both identities use the same I_n : $AI_n = A = I_n A$. The smallest case is $I_1 = [1]$, the 1×1 identity, and $[1][a] = [a]$.

Check: with the column formula of p. 53, $AI_n = [Ae_1 \cdots Ae_n] = [a_1 \cdots a_n] = A$, since $Ae_j = a_j$ (9/3 notes, p. 41). E.g. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} I_2$ has columns (1,3) and (2,4). $\checkmark$

Is $A(BC) = (AB)C$? Yes: matrix multiplication is associative

Let R, T, S be the linear maps of C, B, A (so $R(x) = Cx$, $T(y) = By$, $S(z) = Az$, with sizes such that the products are defined). The notes' hint is

$$(S \circ T \circ R)(x) = S(T(R(x))).$$

Composition of functions is associative: both $S \circ (T \circ R)$ and $(S \circ T) \circ R$ send x to $S(T(R(x)))$, since each one simply applies R , then T , then S . So they are the same map, and therefore they have the same standard matrix. By the definition of the product, the matrix of $S \circ (T \circ R)$ is $A(BC)$ and the matrix of $(S \circ T) \circ R$ is $(AB)C$. Hence $A(BC) = (AB)C$.

Check: for random $2 \times 3, 3 \times 4, 4 \times 2$ matrices, NumPy gives $A(BC) = (AB)C$ to machine precision. $\checkmark$

Why it matters: we may write ABC without brackets, and $A^n = A \cdots A$ (used on p. 56) is unambiguous.

Is $AB = BA$? Usually no

Sizes. If A is $m \times n$ and B is $n \times m$ (so both products exist), then AB is $m \times m$ and BA is $n \times n$. If $m \neq n$ they cannot be equal: they are not even the same size. For example, a 1×2 row times a 2×1 column is 1×1 , but in the other order it is 2×2 .

Even for square matrices ($m = n$) it usually fails. The notes give an example on p. 55 (Section 6): $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ but $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Meaning: AB is "do B , then A " and BA is "do A , then B ". Doing two operations in a different

order usually gives a different result (e.g. “project onto the x_1 -axis, then rotate” versus “rotate, then project”).

4 A formula for the product: column by column (p. 53)

To multiply BA , multiply B by each column of A .

Theorem (formula for the product, p. 53)

If $A = [a_1 \cdots a_n]$ is $m \times n$ and B is $k \times m$, then

$$BA = [Ba_1 \cdots Ba_n].$$

(The notes write “ $S \circ T$ ” under BA as a reminder that BA is the matrix of $S \circ T$.)

Example (p. 53). Here $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $A = \begin{bmatrix} 5 & 0 \\ -1 & 2 \end{bmatrix}$, with columns $a_1 = (5, -1)$ and $a_2 = (0, 2)$. Recall $Bx = x_1(\text{col } 1) + x_2(\text{col } 2)$:

$$\begin{aligned} Ba_1 &= 5 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + (-1) \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \cdot 1 + 2(-1) \\ 5 \cdot 3 + 4(-1) \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \end{bmatrix}, \\ Ba_2 &= 0 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 0 + 2 \cdot 2 \\ 3 \cdot 0 + 4 \cdot 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}. \end{aligned}$$

So

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 11 & 8 \end{bmatrix}.$$

Check that this really is the matrix of “first A , then B ”: take $x = (1, 1)$. $Ax = (5, 1)$, then $B(5, 1)^T = (5 + 2, 15 + 4) = (7, 19)$. Directly, $\begin{bmatrix} 3 & 4 \\ 11 & 8 \end{bmatrix} (1, 1)^T = (7, 19)$. ✓

Proof (end filled in)

Let $T(x) = Ax$ and $S(y) = By$. The standard matrix of $S \circ T$ is, by the recipe of 9/3 notes p. 41 applied to the linear map $S \circ T$,

$$[S(T(e_1)) \cdots S(T(e_n))] = [BT(e_1) \cdots BT(e_n)],$$

because $S(y) = By$ for every y , in particular for $y = T(e_j)$. The matrix for T is $[T(e_1) \cdots T(e_n)] = A$, so $T(e_j) = a_j$, the j -th column of A (also: $Ae_j = a_j$, 9/3 notes, p. 41). Substituting,

$$BA \stackrel{\text{def}}{=} \text{matrix of } S \circ T = [Ba_1 \cdots Ba_n]. \quad \square$$

5 The row–column rule (pp. 54–55)

A single entry of AB is “row i of A times column j of B ”: multiply matching entries and add.

Notation (p. 54)

a_{ij} = the entry of A in row i , column j . We write $A = (a_{ij})$.

Example (p. 54). $A = \begin{bmatrix} 1 & -1 & 0 \\ 7 & 3 & -2 \end{bmatrix} = (a_{ij})$. Then $a_{23} = -2$ (row 2, column 3) and $a_{12} = -1$ (row 1, column 2). “ a_{31} ” makes no sense: A has only 2 rows.

Fill-in: the (1,2) entry of $\begin{bmatrix} 1 & -1 \\ 7 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ is -2

The boxed example on p. 54 circles row 1 of the left matrix and column 2 of the right matrix, and labels it “12” (my reading: the (1,2) entry; no value is written). Multiply matching entries and add:

$$(1)(0) + (-1)(2) = -2.$$

For reference, the whole product is $\begin{bmatrix} 1 & -1 \\ 7 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 4 & 6 \end{bmatrix}$.

Check with the column formula of p. 53: column 2 of the product is $A(0,2)^T = 0(1,7) + 2(-1,3) = (-2,6)$, whose first entry is -2 . ✓

Theorem (row–column rule, p. 54)

If $A = (a_{ij})$ is $m \times n$ and $B = (b_{ij})$ is $n \times k$, then the (i,j) entry of AB is

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

The red underlines in the notes stress that the inner sizes must match: A has n columns and B has n rows, so row i of A and column j of B both have n entries.

Proof (completed)

By the column formula (p. 53, with the roles of the letters swapped), the j -th column of AB is A times the j -th column of B :

$$A \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = b_{1j} \begin{bmatrix} a_{11} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + b_{nj} \begin{bmatrix} a_{1n} \\ \vdots \\ a_{in} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

This step is the definition of matrix–vector multiplication: Ax is $x_1(\text{col } 1) + \cdots + x_n(\text{col } n)$. The (i,j) entry of AB is the i -th entry of this column. Reading off the i -th entry of each vector (the row circled in red in the notes):

$$(AB)_{ij} = b_{1j}a_{i1} + \cdots + b_{nj}a_{in} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}. \quad \square$$

Example (p. 55). The $(2, 1)$ entry of $\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 1 & 3 \end{bmatrix}$ uses row 2 of the left matrix, $(3, -1)$, and column 1 of the right, $(-1, 1)$:

$$3(-1) + (-1)(1) = -4.$$

Check: the full product is $\begin{bmatrix} 1 & 6 \\ -4 & -3 \end{bmatrix}$ (NumPy), with -4 in position $(2, 1)$. ✓

6 Matrices are not numbers (pp. 55–56)

Three rules that hold for real numbers fail for matrices: commutativity, “no zero divisors”, and cancellation.

If A, B are $n \times n$, do we expect $AB = BA$? No

The notes’ counterexample (p. 55), computed with the row–column rule:

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

For instance, the $(1, 2)$ entry of the first product is $(1)(1) + (0)(0) = 1$, while the $(1, 2)$ entry of the second is $(0)(0) + (1)(0) = 0$. So $AB \neq BA$. *Check:* NumPy gives the same two products. ✓

Geometric meaning: $E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ projects onto the x_1 -axis, $N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ sends $(x_1, x_2) \mapsto (x_2, 0)$. “ N then E ” keeps $(x_2, 0)$; “ E then N ” first kills x_2 , so everything goes to 0.

Do we expect $AB = 0 \Rightarrow A = 0$ or $B = 0$? No

The second product above is a counterexample: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq 0$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq 0$, yet $AB = 0$.

Page 56 gives an even more striking one: a nonzero matrix whose *square* is zero,

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

The red circles pick out row 1 of the left factor and column 2 of the right factor: $(0)(1) + (1)(0) = 0$. The other three entries are 0 in the same way. ✓

The side remark on p. 55: “ $\mathbb{Z}/4\mathbb{Z}$ ”. The professor notes that this phenomenon also happens in “clock arithmetic” mod 4, where the numbers are 0, 1, 2, 3 and 4 counts as 0. There, $2 \times 2 = 4 = 0$ although $2 \neq 0$. So “a product of nonzero things can be zero” is not special to matrices, but it never happens for real numbers. (This aside is not needed later in the course.)

Do we expect cancellation, $AB = AC$ and $A \neq 0 \Rightarrow B = C$? No

The notes' counterexample (p. 56): with $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$,

$$A \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix} = A \begin{bmatrix} 5 & 6 \\ 3 & 4 \end{bmatrix},$$

but $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \neq \begin{bmatrix} 5 & 6 \\ 3 & 4 \end{bmatrix}$.

Why: multiplying by A moves row 2 of the right factor up to row 1 and puts zeros in row 2. So A “forgets” the first row of whatever it multiplies, and two matrices that differ only in row 1 give the same product. (With numbers, $ab = ac$, $a \neq 0$ implies $b = c$ because we can divide by a ; a matrix like A cannot be “divided by”.)

Check: NumPy gives $\begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix}$ for both products. ✓

7 Powers of diagonal matrices (pp. 56–57)

To raise a diagonal matrix to a power, raise each diagonal entry to that power.

Example (powers, p. 56). $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ is *diagonal* (all entries off the main diagonal are 0; underlined in red in the notes). By the row–column rule,

$$A^2 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 2 + 0 \cdot 0 & 2 \cdot 0 + 0 \cdot (-1) \\ 0 \cdot 2 + (-1) \cdot 0 & 0 \cdot 0 + (-1)(-1) \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & (-1)^2 \end{bmatrix}.$$

Here A^n means $A \cdot A \cdots A$ (n factors), which is well defined by associativity (p. 52).

Theorem (p. 56)

For $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ and every integer $n \geq 1$, $A^n = \begin{bmatrix} 2^n & 0 \\ 0 & (-1)^n \end{bmatrix}$.

Proof by induction (pp. 56–57, steps explained)

Base case $n = 1$: $A^1 = A = \begin{bmatrix} 2^1 & 0 \\ 0 & (-1)^1 \end{bmatrix}$. True.

Induction hypothesis: assume the formula for $n - 1$: $A^{n-1} = \begin{bmatrix} 2^{n-1} & 0 \\ 0 & (-1)^{n-1} \end{bmatrix}$.

Induction step: $A^n = A A^{n-1}$ (associativity), so by the hypothesis and the row–column rule

$$A^n = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2^{n-1} & 0 \\ 0 & (-1)^{n-1} \end{bmatrix} = \begin{bmatrix} 2 \cdot 2^{n-1} & 0 \\ 0 & (-1)(-1)^{n-1} \end{bmatrix} = \begin{bmatrix} 2^n & 0 \\ 0 & (-1)^n \end{bmatrix}.$$

(The off-diagonal entries are sums of products in which every term contains a 0.) So the formula for $n - 1$ implies the formula for n , and by induction it holds for all $n \geq 1$. □

Check for $n = 3$: $A^3 = \begin{bmatrix} 8 & 0 \\ 0 & -1 \end{bmatrix}$ by direct computation (NumPy), and $2^3 = 8$, $(-1)^3 = -1$. ✓

“Similarly” (p. 57): $\text{diag}(d_1, \dots, d_n)^k = \text{diag}(d_1^k, \dots, d_n^k)$ for $k > 0$, with proof

Write $D = \text{diag}(d_1, \dots, d_n)$ for the diagonal matrix with diagonal entries $d_1, \dots, d_n$ (shorthand, not the notes’ notation). First, the product of two diagonal matrices: by the row–column rule, $(DD')_{ij} = \sum_l d_{il}d'_{lj}$. Every term has $d_{il} = 0$ unless $l = i$, and $d'_{lj} = 0$ unless $l = j$. So the entry is 0 when $i \neq j$, and it is $d_i d'_i$ when $i = j$:

$$\text{diag}(d_1, \dots, d_n) \text{diag}(d'_1, \dots, d'_n) = \text{diag}(d_1 d'_1, \dots, d_n d'_n).$$

Now induct on k exactly as above (base case $k = 1$ is clear):

$$D^k = D D^{k-1} = \text{diag}(d_1, \dots, d_n) \text{diag}(d_1^{k-1}, \dots, d_n^{k-1}) = \text{diag}(d_1^k, \dots, d_n^k). \quad \square$$

Example (p. 57). $\begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ for every choice of signs, since $(\pm 1)^2 = 1$. So the four matrices $\text{diag}(1, 1)$, $\text{diag}(1, -1)$, $\text{diag}(-1, 1)$, $\text{diag}(-1, -1)$ all square to I_2 . Check: NumPy confirms all four. ✓

Meaning: the number 1 has only two square roots, ± 1 , but the matrix I_2 has at least four. For example $\text{diag}(1, -1)$ is the reflection across the x_1 -axis (9/3 notes, p. 36), and reflecting twice returns every vector to where it started.

8 Sums and scalar multiples of matrices (pp. 57–58)

Matrices of the same size can be added and scaled entry by entry, so they behave like vectors.

Sum of maps (p. 57). For two maps $T, S : \mathbb{R}^n \rightarrow \mathbb{R}^m$, define their sum by

$$(T + S)(x) = T(x) + S(x).$$

This motivates the matrix sum on p. 58: if $T(x) = Ax$ and $S(x) = Bx$, then $(T + S)(e_j) = Ae_j + Be_j = a_j + b_j$, so the standard matrix of $T + S$ is obtained by adding A and B column by column, i.e. entry by entry. (This link is our explanation; the notes state the two definitions on consecutive pages.)

Definition (scaling and adding matrices, p. 58)

For a scalar c and an $m \times n$ matrix $A = (a_{ij})$, $cA = (ca_{ij})$: multiply every entry by c .

For two matrices $A = (a_{ij})$, $B = (b_{ij})$ of the same size, $A + B = (a_{ij} + b_{ij})$: add matching entries.

Example (p. 58). $2 \begin{bmatrix} 1 & 3 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ -2 & 10 \end{bmatrix}.$

Fill-in: $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -1 & 5 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 7 \\ 3 & 10 \end{bmatrix}$

The notes stop at $\begin{bmatrix} 1-1 & 2+5 \\ 3+0 & 4+6 \end{bmatrix}$. Finishing the arithmetic: $1 - 1 = 0$, $2 + 5 = 7$, $3 + 0 = 3$,

$4 + 6 = 10$, so

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -1 & 5 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 7 \\ 3 & 10 \end{bmatrix}.$$

Check: NumPy gives the same matrix. ✓

Recall: vectors are things we can add and scale. Since matrices of a fixed size $m \times n$ can be added and scaled in the same way, **matrices can be thought of as vectors**. (Concretely, an $m \times n$ matrix is just a list of mn numbers arranged in a rectangle, and the operations act on each number separately, exactly as for vectors in $\mathbb{R}^{mn}$.)

9 The algebra rules (p. 59)

Sums, scalar multiples and products of matrices obey the familiar rules, except commutativity of the product, so expansions like $(A + B)^2$ must keep AB and BA separate.

“Easy theorem” (p. 59)

For matrices of sizes for which the operations are defined, and scalars r, s (O is the zero matrix):

1. $A + O = O + A = A$
2. $(A + B) + C = A + (B + C)$
3. $r(A + B) = rA + rB$
4. $(r + s)A = rA + sA$
5. $(rs)A = r(sA)$
6. $A(B + C) = AB + AC$
7. $(A + B)C = AC + BC$

Rules 1–5 hold because each entry is just an ordinary number and these rules hold for numbers. Rules 6 and 7, the distributive laws, are the ones that involve the matrix product. The notes draw arrows pointing at them (my reading: to single them out as the new rules, used in the examples below).

Why rules 6 and 7 hold (added proof)

Rule 6. By the column formula (p. 53), column j of $A(B + C)$ is $A(b_j + c_j)$, where b_j, c_j are the j -th columns of B, C . The map $x \mapsto Ax$ is linear (9/3 notes, p. 37), so $A(b_j + c_j) = Ab_j + Ac_j$, which is column j of $AB + AC$.

Rule 7. By the row-column rule, $((A + B)C)_{ij} = \sum_l (a_{il} + b_{il})c_{lj} = \sum_l a_{il}c_{lj} + \sum_l b_{il}c_{lj} = (AC)_{ij} + (BC)_{ij}$. $\square$

Ex: $(A + B)^2 = A^2 + AB + BA + B^2$ and $(A + B)(A - B) = A^2 - AB + BA - B^2$

For $n \times n$ matrices A, B :

$$\begin{aligned}(A + B)^2 &= (A + B)(A + B) \stackrel{(7)}{=} A(A + B) + B(A + B) \stackrel{(6)}{=} A^2 + AB + BA + B^2, \\ (A + B)(A - B) &\stackrel{(7)}{=} A(A - B) + B(A - B) \stackrel{(6)}{=} A^2 - AB + BA - B^2.\end{aligned}$$

The first step uses rule 7 with $C = A + B$ (or $C = A - B$); the second uses rule 6 on each piece. In the second line, $A(-B) = -(AB)$ because the scalar -1 can be pulled out of a product.

The underlined AB and BA cannot be combined into $2AB$, because $AB \neq BA$ in general (p. 55). So the high-school formulas $(a + b)^2 = a^2 + 2ab + b^2$ and $(a + b)(a - b) = a^2 - b^2$ fail for matrices.

Check with $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (the p. 55 pair): $(A + B)^2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}^2 = A^2 + AB + BA + B^2$, but $A^2 + 2AB + B^2 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$. Also $(A + B)(A - B) = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} = A^2 - AB + BA - B^2$, while $A^2 - B^2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. (All computed with NumPy.) ✓

Summary

Fact	Formula
Product = composition	BA is the matrix of $S \circ T$; $(BA)x = B(Ax)$
Sizes	$(k \times m)(m \times n) = k \times n$; inner sizes must agree
Column formula	$BA = [Ba_1 \ \cdots \ Ba_n]$
Row-column rule	$(AB)_{ij} = a_{i1}b_{1j} + \cdots + a_{im}b_{mj}$
Identity	$AI_n = A = I_m A$ for A of size $m \times n$
Associativity	$A(BC) = (AB)C$
Not commutative	$AB \neq BA$ in general; $AB = 0 \not\Rightarrow A = 0$ or $B = 0$; no cancellation
Diagonal powers	$\text{diag}(d_1, \dots, d_n)^k = \text{diag}(d_1^k, \dots, d_n^k)$
Sum, scalar multiple	entry by entry, $A + B$ only for equal sizes
Expansions	$(A + B)^2 = A^2 + AB + BA + B^2$

Corrections to the handwritten notes

Page	As written	Should be	Why
59	$(A + B)(A + B) \stackrel{6}{=} A(A + B) + B(A + B)$	rule 7 (then rule 6 for the next step)	splitting the <i>left</i> factor $A + B$ is $(A + B)C = AC + BC$