Math 201-003 Fall 2019 (Darij Grinberg): homework set 1 with solutions

Exercise 1. **(a)** Write down the augmented matrix corresponding to the following system of linear equations:

$$\begin{cases} 3x - 6y + z = 5 \\ y - 2z = 8 \\ y - 3z = 1 \\ 2z = 5 \end{cases}.$$

(b) Write down the system of linear equations that corresponds to the following augmented matrix:

$$\begin{pmatrix} 1 & 2 & 3 & \mathbf{1} \\ 2 & 3 & 4 & \mathbf{2} \\ 3 & 4 & 5 & \mathbf{3} \end{pmatrix}.$$

(Here, as in class, we are putting the entries of the last column in boldface instead of drawing a vertical bar to the left of them.)

(c) Find the RREF of the matrix in part **(a)**.

(d) Solve the system of linear equations in part **(a)**.

(e) Find the RREF of the matrix in part **(b)**.

(f) Solve the system of linear equations in part **(b)**.

Solution to Exercise 1. **(a)** The augmented matrix is $\begin{pmatrix} 3 & -6 & 1 & \mathbf{5} \\ 0 & 1 & -2 & \mathbf{8} \\ 0 & 1 & -3 & \mathbf{1} \\ 0 & 0 & 2 & \mathbf{5} \end{pmatrix}$.

(b) The system is $\begin{cases} x + 2y + 3z = 1 \\ 2x + 3y + 4z = 2 \\ 3x + 4y + 5z = 3 \end{cases}$.

(c) We shall mark the pivot entries by putting them in boxes. Note that the following is only one way of performing the row-reduction algorithm ([Strickland, Method 6.3]); there are others (since you have freedom in pivoting).

$$\begin{pmatrix} \boxed{3} & -6 & 1 & 5 \\ 0 & 1 & -2 & 8 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 2 & 5 \end{pmatrix}$$

$$\xrightarrow{\text{scale row 1 by } 1/3} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 \\ 0 & 1 & -2 & 8 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 2 & 5 \end{pmatrix}$$

$$\overset{\text{freeze row 1}}{\longrightarrow} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & 1 & -2 & 8 & \\ 0 & 1 & -3 & 1 & \\ 0 & 0 & 2 & 5 & \end{pmatrix}$$

$$\underset{\text{(remember that frozen rows don't count!)}}{\overset{\text{add } -1 \text{ times row 1 to row 2}}{\longrightarrow}} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \\ 0 & 0 & -1 & -7 & \\ 0 & 0 & 2 & 5 & \end{pmatrix}$$

$$\overset{\text{freeze row 1}}{\longrightarrow} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \leftarrow \text{frozen} \\ 0 & 0 & -1 & -7 & \\ 0 & 0 & 2 & 5 & \end{pmatrix}$$

$$\overset{\text{scale row 1 by } -1}{\longrightarrow} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \leftarrow \text{frozen} \\ 0 & 0 & \boxed{1} & 7 & \\ 0 & 0 & 2 & 5 & \end{pmatrix}$$

$$\overset{\text{add } -2 \text{ times row 1 to row 2}}{\longrightarrow} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \leftarrow \text{frozen} \\ 0 & 0 & \boxed{1} & 7 & \\ 0 & 0 & 0 & -9 & \end{pmatrix}$$

$$\overset{\text{freeze row 1}}{\longrightarrow} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \leftarrow \text{frozen} \\ 0 & 0 & \boxed{1} & 7 & \leftarrow \text{frozen} \\ 0 & 0 & 0 & -9 & \end{pmatrix}$$

$$\overset{\text{scale row 1 by } -1/9}{\longrightarrow} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \leftarrow \text{frozen} \\ 0 & 0 & \boxed{1} & 7 & \leftarrow \text{frozen} \\ 0 & 0 & 0 & \boxed{1} & \end{pmatrix}$$

$$\overset{\text{unfreeze last frozen row}}{\longrightarrow} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \leftarrow \text{frozen} \\ 0 & 0 & \boxed{1} & 7 & \\ 0 & 0 & 0 & \boxed{1} & \end{pmatrix}$$

$$\left(\begin{array}{c} \text{before this step, we should have frozen row 1 and} \\ \text{then unfrozen it again, but this changes nothing} \end{array} \right)$$

$$\underset{\longrightarrow}{\text{add } -7 \text{ times row 2 to row 1}} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \leftarrow \text{frozen} \\ 0 & 0 & \boxed{1} & 0 & \\ 0 & 0 & 0 & \boxed{1} & \end{pmatrix}$$

$$\underset{\longrightarrow}{\text{unfreeze last frozen row}} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & -2 & 8 & \\ 0 & 0 & \boxed{1} & 0 & \\ 0 & 0 & 0 & \boxed{1} & \end{pmatrix}$$

$$\underset{\longrightarrow}{\text{add 2 times row 2 to row 1}} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & 0 & 8 & \\ 0 & 0 & \boxed{1} & 0 & \\ 0 & 0 & 0 & \boxed{1} & \end{pmatrix}$$

$$\underset{\longrightarrow}{\text{add } -8 \text{ times row 3 to row 1}} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & 0 & 0 & \\ 0 & 0 & \boxed{1} & 0 & \\ 0 & 0 & 0 & \boxed{1} & \end{pmatrix}$$

$$\underset{\longrightarrow}{\text{unfreeze last frozen row}} \begin{pmatrix} \boxed{1} & -3 & 1/3 & 5/3 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

$$\underset{\longrightarrow}{\text{add 3 times row 2 to row 1}} \begin{pmatrix} \boxed{1} & 0 & 1/3 & 5/3 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

$$\underset{\longrightarrow}{\text{add } -1/3 \text{ times row 3 to row 1}} \begin{pmatrix} \boxed{1} & 0 & 0 & 5/3 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

$$\underset{\longrightarrow}{\text{add } -5/3 \text{ times row 4 to row 1}} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}.$$

This is a matrix in RREF.

(d) The RREF we have just obtained has a pivot in its last column. Thus, the system of linear equations contains an equation saying $0 = 1$, and therefore has no solutions. (See [Strickland, Method 5.4 (b)] for the logic we are using here.)

(e) We shall mark the pivot entries by putting them in boxes. Note that the following is only one way of performing the row-reduction algorithm ([Strickland, Method 6.3]); there are others (since you have freedom in pivoting).

$$\begin{pmatrix} \boxed{1} & 2 & 3 & 1 \\ 2 & 3 & 4 & 2 \\ 3 & 4 & 5 & 3 \end{pmatrix}$$

$$\xrightarrow{\text{add } -2 \text{ times row 1 to row 2}} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 \\ 0 & -1 & -2 & 0 \\ 3 & 4 & 5 & 3 \end{pmatrix}$$

$$\xrightarrow{\text{add } -3 \text{ times row 1 to row 2}} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 \\ 0 & -1 & -2 & 0 \\ 0 & -2 & -4 & 0 \end{pmatrix}$$

$$\xrightarrow{\text{freeze row 1}} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 & \leftarrow \text{frozen} \\ 0 & -1 & -2 & 0 & \\ 0 & -2 & -4 & 0 & \end{pmatrix}$$

$$\xrightarrow{\text{scale row 1 by } -1} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & 2 & 0 & \\ 0 & -2 & -4 & 0 & \end{pmatrix}$$

$$\xrightarrow{\text{add 2 times row 1 to row 2}} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & 2 & 0 & \\ 0 & 0 & 0 & 0 & \end{pmatrix}$$

$$\xrightarrow{\text{freeze row 1}} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & 2 & 0 & \leftarrow \text{frozen} \\ 0 & 0 & 0 & 0 & \end{pmatrix}$$

$$\xrightarrow{\text{unfreeze last frozen row}} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 & \leftarrow \text{frozen} \\ 0 & \boxed{1} & 2 & 0 & \\ 0 & 0 & 0 & 0 & \end{pmatrix}$$

$$\left(\begin{array}{c} \text{we are starting to unfreeze, since we have} \\ \text{obtained a matrix with no nonzero entries} \end{array}\right)$$

$$\xrightarrow{\text{unfreeze last frozen row}} \begin{pmatrix} \boxed{1} & 2 & 3 & 1 \\ 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{\text{add } -2 \text{ times row 2 to row 1}} \begin{pmatrix} \boxed{1} & 0 & -1 & 1 \\ 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This is a matrix in RREF.

(f) We have just row-reduced the matrix that corresponds to the system, and obtained $\begin{pmatrix} \boxed{1} & 0 & -1 & 1 \\ 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Hence, the system is equivalent to the system

$$\begin{cases} x - z = 1 \\ y + 2z = 0 \\ 0 = 0 \end{cases}$$

(where we are labeling the variables by x, y, z). This system has one free variable, namely z (because the third column of the above matrix contains no pivot); the two other variables (that is, x and y) are dependent variables (because the first two columns of the above matrix contain pivots). Thus, its solution is given by

$$x = 1 + z \qquad \text{and} \qquad y = -2z,$$

with z an arbitrary number. (See [Strickland, Method 5.4 (d)] for the logic we have used here.) □

Exercise 2. Consider the two matrices

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

Let $v = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$ be any column vector of size 4.

(a) Compute Cv.
(b) Compute Dv.
(c) Compute CD.
(d) Compute DC.

Solution to Exercise 2. **(a)** From $C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ and $v = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$, we obtain

$$Cv = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1a + 0b + 0c + 0d \\ 1a + 1b + 0c + 0d \\ 1a + 1b + 1c + 0d \\ 1a + 1b + 1c + 1d \end{pmatrix} = \begin{pmatrix} a \\ a + b \\ a + b + c \\ a + b + c + d \end{pmatrix}.$$

(b) From $D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$ and $v = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$, we obtain

$$Dv = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1a + 0b + 0c + 0d \\ (-1)\,a + 1b + 0c + 0d \\ 0a + (-1)\,b + 1c + 0d \\ 0a + 0b + (-1)\,c + 1d \end{pmatrix} = \begin{pmatrix} a \\ b-a \\ c-b \\ d-c \end{pmatrix}.$$

(c) From $C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ and $D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$, we obtain

$$\begin{aligned} CD &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &\qquad \left(\begin{array}{c}\text{by directly applying the definition of the} \\ \text{product of two matrices, and simplifying}\end{array}\right) \\ &= I_4. \end{aligned}$$

(d) Similarly, a computation shows that $DC = I_4$.
[*Remark:* The answers to parts **(a)** and **(b)** can be restated as follows:

- The vector Cv is obtained from v by **cumulating** the entries of v from top to bottom.
- The vector Dv is obtained from v by taking **differences** between consecutive entries of v.

This is why we called the matrices C and D. More usefully, this gives an alternative explanation for why $CD = I_4$ and $DC = I_4$, without having to actually multiply out the matrices. Indeed, any matrix A is uniquely determined by its products with all column vectors; in other words, if A_1 and A_2 are two $n \times m$-matrices such that

$$A_1 v = A_2 v \qquad \text{for all } v \in \mathbb{R}^m, \tag{1}$$

then $A_1 = A_2$. [1] Thus, in order to prove that $CD = I_4$, it suffices to show that $CDv = I_4 v$ for all $v \in \mathbb{R}^4$. But this is easy: Due to what we have just said about C and D, we see that

[1]This is not hard to prove (we will do this in class). Just fix any $j \in \{1, 2, \ldots, m\}$, and apply the assumption (1) to $v = (0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)^T$ (where the 1 is in the j-th position of the vector); the result will say that the j-th column of A_1 equals the j-th column of A_2. Since this holds for all $j \in \{1, 2, \ldots, m\}$, we can thus conclude that $A_1 = A_2$.

- the vector Dv is obtained from v by taking differences between consecutive entries of v, and
- the vector CDv is obtained from Dv by cumulating the entries of v from top to bottom.

But this procedure (in which we first take differences between consecutive entries, and then cumulate these differences) just returns our original vector v, which of course is the same as $I_4 v$. Thus, $CDv = I_4 v$. Since this holds for all $v \in \mathbb{R}^4$, we thus conclude that $CD = I_4$. Likewise, $DC = I_4$, because cumulating and then taking consecutive difference is another way to get the original vector back.] $\square$

Definition 0.1. Let A be any square matrix. Then, the *trace* of A is defined to be the sum of all diagonal entries of A. In other words, it is defined to be $A_{1,1} + A_{2,2} + \cdots + A_{n,n}$, where $n \in \mathbb{N}$ is such that A is an $n \times n$-matrix.

The trace of A is denoted by $\operatorname{Tr} A$.

For example, here is what the traces of 2×2-matrices and 3×3-matrices are:

$$\operatorname{Tr}\begin{pmatrix} a & b \\ a' & b' \end{pmatrix} = a + b'; \qquad \operatorname{Tr}\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = a + b' + c''.$$

Exercise 3. **(a)** Prove that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ for any two 2×2-matrices A and B.

(b) Prove that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ for any two 3×3-matrices A and B.

[**Hint:** In part **(a)**, you can set $A = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$ and $B = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}$ and compute both $\operatorname{Tr}(AB)$ and $\operatorname{Tr}(BA)$ and check their equality. Likewise for part **(b)**. Alternatively, you can show that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ for any $n \times n$-matrices A and B of any size n. In the latter case, you will want to be familiar with the basics of summation signs [lina, §2.9].)

First solution to Exercise 3. **(a)** Let A and B be two 2×2-matrices. Write A and B as $A = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$ and $B = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}$. Thus,

$$AB = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}\begin{pmatrix} x & y \\ x' & y' \end{pmatrix} = \begin{pmatrix} ax + bx' & ay + by' \\ a'x + b'x' & a'y + b'y' \end{pmatrix},$$

so that the definition of a trace yields

$$\operatorname{Tr}(AB) = (ax + bx') + (a'y + b'y'). \tag{2}$$

Also,

$$BA = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}\begin{pmatrix} a & b \\ a' & b' \end{pmatrix} = \begin{pmatrix} xa + ya' & xb + yb' \\ x'a + y'a' & x'b + y'b' \end{pmatrix},$$

so that the definition of a trace yields

$$\operatorname{Tr}(BA) = (xa + ya') + (x'b + y'b'). \tag{3}$$

We need to prove that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. In view of the equalities (2) and (3), this rewrites as

$$(ax + bx') + (a'y + b'y') = (xa + ya') + (x'b + y'b').$$

But verifying this is a matter of straightforward computation:

$$\begin{aligned}(ax + bx') + (a'y + b'y') &= \underbrace{ax}_{=xa} + \underbrace{bx'}_{=x'b} + \underbrace{a'y}_{=ya'} + \underbrace{b'y'}_{=y'b'} \\ &= xa + x'b + ya' + y'b' = (xa + ya') + (x'b + y'b').\end{aligned}$$

Thus, $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ is proven. This solves Exercise 3 **(a)**.

Before we solve part **(b)**, let us make a simple but useful observation: We did not need to compute the two off-diagonal entries of AB and BA, since the trace of a matrix only depends on its diagonal entries. This will save us some work in our solution to part **(b)**:

(b) Let A and B be two 3×3-matrices. Write A and B as $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$ and $B = \begin{pmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{pmatrix}$. Thus,

$$\begin{aligned}AB &= \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \begin{pmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{pmatrix} \\ &= \begin{pmatrix} ax + bx' + cx'' & * & * \\ * & a'y + b'y' + c'y'' & * \\ * & * & a''z + b''z' + c''z'' \end{pmatrix},\end{aligned}$$

where the "$*$" symbols stand for entries that we don't want to compute. Thus, the definition of a trace yields

$$\operatorname{Tr}(AB) = (ax + bx' + cx'') + (a'y + b'y' + c'y'') + (a''z + b''z' + c''z''). \tag{4}$$

A similar computation of BA shows that

$$\operatorname{Tr}(BA) = (xa + ya' + za'') + (x'b + y'b' + z'b'') + (x''c + y''c' + z''c''). \tag{5}$$

We need to prove that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. In view of the equalities (4) and (5), this rewrites as

$$(ax + bx' + cx'') + (a'y + b'y' + c'y'') + (a''z + b''z' + c''z'')$$

$$= \left(xa + ya' + za''\right) + \left(x'b + y'b' + z'b''\right) + \left(x''c + y''c' + z''c''\right).$$

But verifying this is a matter of straightforward computation:

$$\begin{aligned}
&\left(ax + bx' + cx''\right) + \left(a'y + b'y' + c'y''\right) + \left(a''z + b''z' + c''z''\right) \\
&= \underbrace{ax}_{=xa} + \underbrace{bx'}_{=x'b} + \underbrace{cx''}_{=x''c} + \underbrace{a'y}_{=ya'} + \underbrace{b'y'}_{=y'b'} + \underbrace{c'y''}_{=y''c'} + \underbrace{a''z}_{=za''} + \underbrace{b''z'}_{=z'b''} + \underbrace{c''z''}_{=z''c''} \\
&= xa + x'b + x''c + ya' + y'b' + y''c' + za'' + z'b'' + z''c'' \\
&= \left(xa + ya' + za''\right) + \left(x'b + y'b' + z'b''\right) + \left(x''c + y''c' + z''c''\right).
\end{aligned}$$

Thus, $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ is proven. This solves Exercise 3 **(b)**. □

Second solution to Exercise 3. As the Hint suggests, we can kill two birds with one stone by proving that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ holds for any two $n \times n$-matrices A and B for any $n \in \mathbb{N}$. This requires some familiarity with summation signs; see [lina, §2.9] for an introduction.

Let us show this proof. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices. We must show that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. Recall that we are using the notation $C_{i,j}$ for the (i,j)-th entry[2] of a matrix C. Now, [lina, Proposition 2.28] (applied to $m = n$ and $p = n$) yields

$$(AB)_{i,j} = \sum_{k=1}^{n} A_{i,k} B_{k,j} \qquad (6)$$

for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, n\}$. But the definition of a trace shows that

$$\begin{aligned}
\operatorname{Tr}(AB) &= (AB)_{1,1} + (AB)_{2,2} + \cdots + (AB)_{n,n} = \sum_{\ell=1}^{n} \underbrace{(AB)_{\ell,\ell}}_{\substack{= \sum_{k=1}^{n} A_{\ell,k} B_{k,\ell} \\ \text{(by (6), applied to } i=\ell \text{ and } j=\ell)}} \\
&= \sum_{\ell=1}^{n} \sum_{k=1}^{n} A_{\ell,k} B_{k,\ell}. \qquad (7)
\end{aligned}$$

The same argument (but with the roles of A and B interchanged) shows that

$$\operatorname{Tr}(BA) = \sum_{\ell=1}^{n} \sum_{k=1}^{n} B_{\ell,k} A_{k,\ell}. \qquad (8)$$

We want to prove that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. In view of (7) and (8), this rewrites as

$$\sum_{\ell=1}^{n} \sum_{k=1}^{n} A_{\ell,k} B_{k,\ell} = \sum_{\ell=1}^{n} \sum_{k=1}^{n} B_{\ell,k} A_{k,\ell}.$$

[2] that is, the entry in row i and column j

Now, to prove this equality, we need to perform several manipulations of sums. First of all, we rename the two summation indexes ℓ and k in the sum $\sum\limits_{\ell=1}^{n}\sum\limits_{k=1}^{n} A_{\ell,k}B_{k,\ell}$ as k and ℓ, respectively. We thus obtain

$$\sum_{\ell=1}^{n}\sum_{k=1}^{n} A_{\ell,k}B_{k,\ell} = \sum_{k=1}^{n}\sum_{\ell=1}^{n} \underbrace{A_{k,\ell}B_{\ell,k}}_{\substack{=B_{\ell,k}A_{k,\ell} \\ \text{(since multiplication of} \\ \text{numbers is commutative)}}} = \sum_{k=1}^{n}\sum_{\ell=1}^{n} B_{\ell,k}A_{k,\ell}.$$

Next, we interchange the two summation signs; to be more precise, we apply [lina, Proposition 2.31] to $m = n$, $p = n$ and $a_{k,\ell} = B_{\ell,k}A_{k,\ell}$. Thus, we obtain

$$\sum_{k=1}^{n}\sum_{\ell=1}^{n} B_{\ell,k}A_{k,\ell} = \sum_{\ell=1}^{n}\sum_{k=1}^{n} B_{\ell,k}A_{k,\ell}.$$

Combining what we have proven, we obtain

$$\sum_{\ell=1}^{n}\sum_{k=1}^{n} A_{\ell,k}B_{k,\ell} = \sum_{k=1}^{n}\sum_{\ell=1}^{n} B_{\ell,k}A_{k,\ell} = \sum_{\ell=1}^{n}\sum_{k=1}^{n} B_{\ell,k}A_{k,\ell}.$$

In view of (7) and (8), this rewrites as $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. This is what we wanted to show. Thus, Exercise 3 is solved again. $\square$

Exercise 4. Let $n \in \mathbb{N}$. For each $k \in \{1, 2, \ldots, n\}$, we let e_k denote the $n \times 1$-matrix (i.e., the column vector of size n) whose k-th entry is 1 and whose all other entries are 0.

(For example: If $n = 3$, then $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.)

(a) If $n = 3$ and if $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$ is any 2×3-matrix, then compute Ae_1, Ae_2 and Ae_3.

(b) Express Ae_k in very simple terms (no matrix multiplication involved!) for every $m \times n$-matrix A (of arbitrary dimensions m and n).

(c) If $n = 2$ and if $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$ is any 2×3-matrix, then compute $e_1^T A$ and $e_2^T A$. (Note: e_k^T means $(e_k)^T$, that is, the transpose of e_k. This is a row vector.)

(d) Express $e_k^T A$ in very simple terms (no matrix multiplication involved!) for every $n \times m$-matrix A (of arbitrary dimensions n and m).

[In parts **(b)** and **(d)**, you don't need to prove your answer.]

Solution to Exercise 4. **(a)** If $n = 3$ and $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$, then

$$Ae_1 = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a \cdot 1 + b \cdot 0 + c \cdot 0 \\ a' \cdot 1 + b' \cdot 0 + c' \cdot 0 \end{pmatrix} = \begin{pmatrix} a \\ a' \end{pmatrix}$$

and similarly

$$Ae_2 = \begin{pmatrix} b \\ b' \end{pmatrix} \qquad \text{and} \qquad Ae_3 = \begin{pmatrix} c \\ c' \end{pmatrix}.$$

(b) If A is any $m \times n$-matrix, and if $k \in \{1, 2, \ldots, n\}$, then $Ae_k = \operatorname{col}_k A$. (Here, we are using the notations of [lina, Definition 2.17 **(b)**]; thus, $\operatorname{col}_k A$ denotes the k-th column of A.)

The easiest way to prove this is using [lina, Proposition 2.19]. Indeed, let A be any $m \times n$-matrix, and let $k \in \{1, 2, \ldots, n\}$. Note that the identity matrix I_n satisfies $\operatorname{col}_k (I_n) = e_k$ (since both $\operatorname{col}_k (I_n)$ and e_k are column vectors of size n, with their k-th entries equal to 1 and all other entries equal to 0). But [lina, Proposition 2.19 **(d)**] (applied to m, n, n, I_n and k instead of n, m, p, B and j) shows that $\operatorname{col}_k (AI_n) = A \cdot \operatorname{col}_k (I_n)$. Since $AI_n = A$ and $\operatorname{col}_k (I_n) = e_k$, this rewrites as $\operatorname{col}_k A = A \cdot e_k = Ae_k$. In other words, $Ae_k = \operatorname{col}_k A$, which is precisely what we wanted to prove.

(c) If $n = 2$ and $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$, then

$$e_1^T A = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = \begin{pmatrix} 1a + 0a' & 1b + 0b' & 1c + 0c' \end{pmatrix} = \begin{pmatrix} a & b & c \end{pmatrix}$$

and similarly

$$e_2^T A = \begin{pmatrix} a' & b' & c' \end{pmatrix}.$$

(d) If A is any $n \times m$-matrix, and if $k \in \{1, 2, \ldots, n\}$, then $e_k^T A = \operatorname{row}_k A$. (Here, we are using the notations of [lina, Definition 2.17 **(a)**]; thus, $\operatorname{row}_k A$ denotes the k-th row of A.)

The easiest way to prove this is using [lina, Proposition 2.19]. Indeed, let A be any $n \times m$-matrix, and let $k \in \{1, 2, \ldots, n\}$. Note that the identity matrix I_n satisfies $\operatorname{row}_k (I_n) = e_k^T$ (since both $\operatorname{row}_k (I_n)$ and e_k^T are row vectors of size n, with their k-th entries equal to 1 and all other entries equal to 0). But [lina, Proposition 2.19 **(c)**] (applied to n, n, m, I_n, A and k instead of n, m, p, A, B and i) shows that $\operatorname{row}_k (I_n A) = (\operatorname{row}_k (I_n)) \cdot A$. Since $I_n A = A$ and $\operatorname{row}_k (I_n) = e_k^T$, this rewrites as $\operatorname{row}_k A = e_k^T \cdot A = e_k^T A$. In other words, $e_k^T A = \operatorname{row}_k A$, which is precisely what we wanted to prove. □

Exercise 5. Which of the following matrices are in RREF? (See [Strickland, Definition 5.1] for the definition of an RREF.)

(a) $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.
(b) $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.
(c) $\begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

(d) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.
(e) $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.
(f) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

(g) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. **(h)** $\begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$ **(i)** $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

(j) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Solution to Exercise 5. See [Strickland, Definition 5.1] for the properties RREF0, RREF1, RREF2 and RREF3 that we will refer to in the following.

(a) This matrix is not in RREF, since it violates property RREF3. (Namely, the pivot in the second row has a nonzero entry in its column.)

(b) This matrix is in RREF.

(c) This matrix is not in RREF, since it violates property RREF3. (Namely, the pivot in the second row has a nonzero entry in its column.)

(d) This matrix is not in RREF, since it violates property RREF1 (because two of its pivots are not equal to 1).

(e) This matrix is not in RREF, since it violates property RREF3. (Namely, the pivot in the first row has a nonzero entry in its column.) It also violates property RREF2 (since the pivot in the second row is not further right than the pivot in the first row).

(f) This matrix is in RREF.

(g) This matrix is not in RREF, since it violates property RREF0. (Indeed, its second row is zero, but its third row is nonzero, so it has a zero row that does not come after all the nonzero rows.)

(h) This matrix is not in RREF, since it violates property RREF1. (It also violates RREF2 and RREF3.)

(i) This matrix is not in RREF, since it violates property RREF3. (Namely, the pivot in the second row has a nonzero entry in its column.)

(j) This matrix is in RREF. (It is a zero matrix, and any zero matrix is in RREF.) □

Exercise 6. **(a)** Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Show that the matrices 2×2-matrices B that satisfy $AB = BA$ are precisely the matrices of the form $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$ for $a, b \in \mathbb{R}$.

[**Hint:** Let $B = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}$ be any 2×2-matrix. Your goal is to prove that $AB = BA$ holds if and only if $x = y'$ and $x' = 0$.]

(b) Let $A' = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Characterize the 2×2-matrices B that satisfy $A'B = BA'$.

[That is, find a similar characterization to the one given in part **(a)**. Show your reasoning!]

Solution to Exercise 6. **(a)** Let $B = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}$ be any 2×2-matrix.

Multiplying the equalities $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}$, we obtain

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ x' & y' \end{pmatrix} = \begin{pmatrix} 1x + 1x' & 1y + 1y' \\ 0x + 1x' & 0y + 1y' \end{pmatrix} = \begin{pmatrix} x + x' & y + y' \\ x' & y' \end{pmatrix}.$$

Similarly,

$$BA = \begin{pmatrix} x & x + y \\ x' & x' + y' \end{pmatrix}.$$

Now, we have the following chain of logical equivalences:

$$\begin{aligned}
&(AB = BA) \\
\Longleftrightarrow\ & \left(\begin{pmatrix} x + x' & y + y' \\ x' & y' \end{pmatrix} = \begin{pmatrix} x & x + y \\ x' & x' + y' \end{pmatrix} \right) \\
& \qquad \left(\text{since } AB = \begin{pmatrix} x + x' & y + y' \\ x' & y' \end{pmatrix} \text{ and } BA = \begin{pmatrix} x & x + y \\ x' & x' + y' \end{pmatrix} \right) \\
\Longleftrightarrow\ & (x + x' = x \text{ and } y + y' = x + y \text{ and } x' = x' \text{ and } y' = x' + y') \\
& \qquad \left(\begin{array}{c} \text{because two matrices are equal if and only if} \\ \text{each entry of the former equals the} \\ \text{respective entry of the latter} \end{array} \right) \\
\Longleftrightarrow\ & (x' = 0 \text{ and } x = y') \\
& \qquad \left(\begin{array}{c} \text{because both equations } x + x' = x \text{ and } y' = x' + y' \\ \text{are equivalent to } x' = 0, \\ \text{whereas the equation } x' = x' \text{ is a tautology (i.e., always true),} \\ \text{while the equation } y + y' = x + y \text{ is equivalent to } y' = x \end{array} \right) \\
\Longleftrightarrow\ & (x = y' \text{ and } x' = 0) \\
\Longleftrightarrow\ & (\text{the diagonal entries of } B \text{ are equal, whereas the } (2,1)\text{-th entry of } B \text{ equals } 0) \\
& \qquad \left(\begin{array}{c} \text{since } x \text{ and } y' \text{ are the diagonal entries of } B, \\ \text{whereas } x' \text{ is the } (2,1)\text{-th entry of } B \end{array} \right) \\
\Longleftrightarrow\ & \left(B = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \text{ for some } a, b \in \mathbb{R} \right).
\end{aligned}$$

This solves part **(a)** of the exercise.

(b) Let $B = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}$ be any 2×2-matrix.

Multiplying the equalities $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} x & y \\ x' & y' \end{pmatrix}$, we obtain

$$AB = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ x' & y' \end{pmatrix} = \begin{pmatrix} 1x + 1x' & 1y + 1y' \\ 1x + 1x' & 1y + 1y' \end{pmatrix} = \begin{pmatrix} x + x' & y + y' \\ x + x' & y + y' \end{pmatrix}.$$

Similarly,

$$BA = \begin{pmatrix} x+y & x+y \\ x'+y' & x'+y' \end{pmatrix}.$$

Now, we have the following chain of logical equivalences:

$$\begin{aligned}
&(AB = BA) \\
\iff & \left(\begin{pmatrix} x+x' & y+y' \\ x+x' & y+y' \end{pmatrix} = \begin{pmatrix} x+y & x+y \\ x'+y' & x'+y' \end{pmatrix} \right) \\
& \qquad \left(\text{since } AB = \begin{pmatrix} x+x' & y+y' \\ x+x' & y+y' \end{pmatrix} \text{ and } BA = \begin{pmatrix} x+y & x+y \\ x'+y' & x'+y' \end{pmatrix} \right) \\
\iff & \left(x+x' = x+y \text{ and } y+y' = x+y \text{ and } x+x' = x'+y' \text{ and } y+y' = x'+y' \right) \\
& \qquad \left(\begin{array}{c} \text{because two matrices are equal if and only if} \\ \text{each entry of the former equals the} \\ \text{respective entry of the latter} \end{array} \right) \\
\iff & \left(x' = y \text{ and } y' = x \right) \\
& \qquad \left(\begin{array}{c} \text{because both equations } x+x' = x+y \text{ and } y+y' = x'+y' \\ \text{are equivalent to } x' = y, \\ \text{whereas both equations } y+y' = x+y \text{ and } x+x' = x'+y' \\ \text{are equivalent to } y' = x \end{array} \right) \\
\iff & \left(\text{the off-diagonal entries of } B \text{ are equal, and the diagonal entries of } B \text{ are equal} \right) \\
& \qquad \left(\begin{array}{c} \text{since } x' \text{ and } y \text{ are the off-diagonal entries of } B, \\ \text{whereas } x \text{ and } y' \text{ are the diagonal entries of } B \end{array} \right) \\
\iff & \left(B = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \text{ for some } a, b \in \mathbb{R} \right).
\end{aligned}$$

[*Remark:* We could have approached both parts **(a)** and **(b)** more systematically: For example, in part **(a)**, instead of manually transforming the equations

$$x + x' = x \text{ and } y + y' = x + y \text{ and } x' = x' \text{ and } y' = x' + y'$$

into simpler forms, we could have treated them as a system of linear equations

$$\begin{cases} x + x' = x \\ y + y' = x + y \\ x' = x' \\ y' = x' + y' \end{cases} \qquad \text{in four unknowns } x, y, x', y',$$

and solved this system using Gaussian elimination.] □

References

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

[Strickland] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 28 September 2013.
http://neil-strickland.staff.shef.ac.uk/courses/MAS201/