Math 201-003 Fall 2019 (Darij Grinberg): homework set 2 with solutions

Exercise 1. Let $n \in \mathbb{N}$. For any n numbers $d_1, d_2, \ldots, d_n \in \mathbb{R}$, we let $\operatorname{diag}(d_1, d_2, \ldots, d_n)$ denote the $n \times n$-matrix whose diagonal entries are $d_1, d_2, \ldots, d_n$ (in this order from top to bottom), while its off-diagonal entries are all 0. In other words,

$$\operatorname{diag}(d_1, d_2, \ldots, d_n) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}.$$

This is called the *diagonal matrix with diagonal entries* $d_1, d_2, \ldots, d_n$.

(a) Given any $d_1, d_2, d_3 \in \mathbb{R}$ and any 3×3-matrix $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$, compute $\operatorname{diag}(d_1, d_2, d_3) \cdot A$.

(b) State a rule for computing $\operatorname{diag}(d_1, d_2, \ldots, d_n) \cdot A$, where $d_1, d_2, \ldots, d_n \in \mathbb{R}$ and where A is any $n \times m$-matrix (with m being any nonnegative integer).

(c) State a rule for computing $A \cdot \operatorname{diag}(d_1, d_2, \ldots, d_n)$, where $d_1, d_2, \ldots, d_n \in \mathbb{R}$ and where A is any $m \times n$-matrix (with m being any nonnegative integer).

[You don't have to prove these rules.]

Solution to Exercise 1. **(a)** We have

$$\operatorname{diag}(d_1, d_2, d_3) = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \qquad \text{(by the definition of } \operatorname{diag}(d_1, d_2, d_3))$$

and

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

Multiplying these equalities, we obtain

$$\begin{aligned} \operatorname{diag}(d_1, d_2, d_3) \cdot A &= \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \\ &= \begin{pmatrix} d_1 a + 0a' + 0a'' & d_1 b + 0b' + 0b'' & d_1 c + 0c' + 0c'' \\ 0a + d_2 a' + 0a'' & 0b + d_2 b' + 0b'' & 0c + d_2 c' + 0c'' \\ 0a + 0a' + d_3 a'' & 0b + 0b' + d_3 b'' & 0c + 0c' + d_3 c'' \end{pmatrix} \\ &= \begin{pmatrix} d_1 a & d_1 b & d_1 c \\ d_2 a' & d_2 b' & d_2 c' \\ d_3 a'' & d_3 b'' & d_3 c'' \end{pmatrix}. \end{aligned}$$

(b) Let $d_1, d_2, \ldots, d_n \in \mathbb{R}$. Let $m \in \mathbb{N}$. Let A be any $n \times m$-matrix. Then, the matrix $\operatorname{diag}(d_1, d_2, \ldots, d_n) \cdot A$ is obtained from A by scaling the i-th row by d_i for each $i \in \{1, 2, \ldots, n\}$. In other words,

$$(\operatorname{diag}(d_1, d_2, \ldots, d_n) \cdot A)_{i,j} = d_i A_{i,j} \qquad (1)$$

for each $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$.

[*Proof of (1):* Set $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$. Thus,

$$D_{i,k} = 0 \qquad \text{for any two distinct elements } i \text{ and } k \text{ of } \{1, 2, \ldots, n\}, \qquad (2)$$

and

$$D_{i,i} = d_i \qquad \text{for each } i \in \{1, 2, \ldots, n\}. \qquad (3)$$

But the definition of the product of two matrices yields

$$DA = \left(D_{i,1}A_{1,j} + D_{i,2}A_{2,j} + \cdots + D_{i,n}A_{n,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}. \qquad (4)$$

Hence,

$$(DA)_{i,j} = D_{i,1}A_{1,j} + D_{i,2}A_{2,j} + \cdots + D_{i,n}A_{n,j} \qquad (5)$$

for each $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$.

Fix $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$. The sum $D_{i,1}A_{1,j} + D_{i,2}A_{2,j} + \cdots + D_{i,n}A_{n,j}$ has n addends. One of these n addends (to be specific: the i-th one) is

$$\underbrace{D_{i,i}}_{\substack{=d_i \\ \text{(by (3))}}} A_{i,j} = d_i A_{i,j},$$

whereas all the other addends are 0 (because each of them has the form $D_{i,k}A_{k,j}$ for some $k \in \{1, 2, \ldots, n\}$ distinct from i, and thus rewrites as $\underbrace{D_{i,k}}_{\substack{=0 \\ \text{(by (2))}}} A_{k,j} = 0$). Hence, this sum equals $d_i A_{i,j}$. In other words,

$$D_{i,1}A_{1,j} + D_{i,2}A_{2,j} + \cdots + D_{i,n}A_{n,j} = d_i A_{i,j}.$$

Hence, (5) rewrites as $(DA)_{i,j} = d_i A_{i,j}$. In view of $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$, this further rewrites as

$$(\operatorname{diag}(d_1, d_2, \ldots, d_n) \cdot A)_{i,j} = d_i A_{i,j}.$$

This proves (1).]

(c) Let $d_1, d_2, \ldots, d_n \in \mathbb{R}$. Let $m \in \mathbb{N}$. Let A be any $m \times n$-matrix. Then, the matrix $A \cdot \operatorname{diag}(d_1, d_2, \ldots, d_n)$ is obtained from A by scaling the j-th column by d_j for each $j \in \{1, 2, \ldots, n\}$. In other words,

$$(A \cdot \operatorname{diag}(d_1, d_2, \ldots, d_n))_{i,j} = d_j A_{i,j}$$

for each $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$.

[The proof of this is similar to the proof of (1); we omit it.] □

We recall the definition of linear combinations ([Strickland, Definition 7.1]):

Definition 0.1. Let $v_1, v_2, \ldots, v_k$ be some vectors in $\mathbb{R}^n$. Then, a *linear combination* of $v_1, v_2, \ldots, v_k$ means a vector that can be written in the form

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k \qquad \text{for some } \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}.$$

The numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$ in this definition are called *coefficients* of this linear combination; but they are not necessarily unique (in fact, they are unique if $v_1, v_2, \ldots, v_k$ are linearly independent; we shall see this later in class). For example, if $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, then the vector $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$ is a linear combination of v_1 and v_2, and can be written as $\lambda_1 v_1 + \lambda_2 v_2$ in many ways:

$$\begin{pmatrix} 3 \\ 3 \end{pmatrix} = 1v_1 + 1v_2 = 3v_1 + 0v_2 = 0v_1 + \frac{3}{2}v_2 = 2v_1 + \frac{1}{2}v_2 = \cdots .$$

We recall the definition of linear independence ([Strickland, §8]):

Definition 0.2. Let $v_1, v_2, \ldots, v_k$ be some vectors in $\mathbb{R}^n$.

(a) A *relation* (more precisely: *linear relation*) between $v_1, v_2, \ldots, v_k$ means a choice of reals $\lambda_1, \lambda_2, \ldots, \lambda_k$ satisfying

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0.$$

(Here, 0 denotes the zero vector $0_{n \times 1} \in \mathbb{R}^n$.)

(b) The *trivial relation* between $v_1, v_2, \ldots, v_k$ is the relation obtained by choosing $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. Clearly, $v_1, v_2, \ldots, v_k$ always have this trivial relation.

(c) We say that the vectors $v_1, v_2, \ldots, v_k$ (or, more precisely, the list $(v_1, v_2, \ldots, v_k)$ of these vectors) are *independent* (more precisely: *linearly independent*) if the only relation between $v_1, v_2, \ldots, v_k$ is the trivial relation. Otherwise, we say that these vectors are *dependent*.

For example, if we set

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \text{and} \qquad v_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

then the vectors v_1, v_2, v_3 are linearly dependent, since they have the nontrivial relation $1v_1 + (-2)v_2 + 1v_3 = 0$.

Exercise 2. Let

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \qquad \text{and} \qquad v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \qquad \text{and} \qquad v_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

(a) Are v_1, v_2, v_3 dependent? (If yes, show a nontrivial relation between them.)

(b) Is $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ a linear combination of v_1, v_2, v_3 ? (If yes, provide a choice of coefficients $\lambda_1, \lambda_2, \lambda_3$ that demonstrate it.)

(c) Is $\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$ a linear combination of v_1, v_2, v_3 ? (If yes, provide a choice of coefficients $\lambda_1, \lambda_2, \lambda_3$ that demonstrate it.)

Solution to Exercise 2. **(a)** The vectors v_1, v_2, v_3 are dependent.

Proof. It is easy to see that $1v_1 + 1v_2 + 1v_3 = 0$, which is a nontrivial relation.

(b) No.

Proof. Each of the three vectors v_1, v_2, v_3 has the property that the sum of its coordinates is 0. Thus, every combination of v_1, v_2, v_3 has this property as well[1]. But the vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ does not have this property. Thus, the latter vector is not a combination of v_1, v_2, v_3.

(c) Yes.

Proof. We have $\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = v_3 + 3v_1 = 3v_1 + 0v_2 + 1v_3$.

[*Remark:* The above proofs rely on pattern-spotting and educated guesses, which require a bit of ingenuity or experience to find. But you can just as well solve the exercise by following algorithms. Indeed, [Strickland, Method 8.8] is a surefire way to answer questions like part **(a)** of this exercise, whereas [Strickland, Method 7.6] can be used to solve parts **(b)** and **(c)**.] □

Clearly, rearranging a list of vectors does not change the set of its linear combinations: In fact, the rearranged vectors will have the same linear combinations as the original vectors; only the coefficients will change their order. For example, the linear combinations of v_1, v_2, v_3 are the same as the linear combinations of v_3, v_1, v_2, since $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \lambda_3 v_3 + \lambda_1 v_1 + \lambda_2 v_2$.

Rearranging a list of vectors also does not change its linear dependence or independence: In fact, any linear relation between the rearranged vectors is a rearranged linear relation between the original vectors, and vice versa. Thus, v_1, v_2, v_3 are dependent if and only if v_3, v_1, v_2 are dependent.

What happens to lists of vectors when we duplicate a vector? The following exercise answers this question (at least when it's the last vector that is duplicated):

Exercise 3. Let $v_1, v_2, \ldots, v_k$ be some vectors in $\mathbb{R}^n$. Assume that $v_{k-1} = v_k$. Justify the following (i.e., sketch the proofs):

[1] Make sure you understand why! (Also make sure you understand why a property like "one of the coordinates is 0" would **not** be transferred from v_1, v_2, v_3 to all their combinations.

(a) The vectors $v_1, v_2, \ldots, v_k$ are linearly dependent.

(b) The linear combinations of $v_1, v_2, \ldots, v_k$ are the same vectors as the linear combinations of $v_1, v_2, \ldots, v_{k-1}$.

Solution to Exercise 3. **(a)** We have

$$0v_1 + 0v_2 + \cdots + 0v_{k-2} + 1v_{k-1} + (-1)\, v_k = v_{k-1} - v_k = 0$$

(since $v_{k-1} = v_k$). Thus, the relation

$$0v_1 + 0v_2 + \cdots + 0v_{k-2} + 1v_{k-1} + (-1)\, v_k = 0$$

between $v_1, v_2, \ldots, v_k$ holds. This relation is clearly nontrivial, since $1 \neq 0$. Hence, the vectors $v_1, v_2, \ldots, v_k$ are linearly dependent. This solves part **(a)** of the exercise.

(b) We need to show the following two claims:

Claim 1: Every linear combination of $v_1, v_2, \ldots, v_k$ is a linear combination of $v_1, v_2, \ldots, v_{k-1}$.

Claim 2: Every linear combination of $v_1, v_2, \ldots, v_{k-1}$ is a linear combination of $v_1, v_2, \ldots, v_k$.

[*Proof of Claim 1:* Let w be a linear combination of $v_1, v_2, \ldots, v_k$. We must show that w is a linear combination of $v_1, v_2, \ldots, v_{k-1}$.

We know that w is a linear combination of $v_1, v_2, \ldots, v_k$. In other words, there exist numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$ such that $w = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k$. Thus,

$$\begin{aligned} w &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-2} v_{k-2} + \lambda_{k-1} v_{k-1} + \lambda_k \underbrace{v_k}_{= v_{k-1}} \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-2} v_{k-2} + \underbrace{\lambda_{k-1} v_{k-1} + \lambda_k v_{k-1}}_{= (\lambda_{k-1} + \lambda_k) v_{k-1}} \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-2} v_{k-2} + (\lambda_{k-1} + \lambda_k)\, v_{k-1}. \end{aligned}$$

This shows that w is a linear combination of $v_1, v_2, \ldots, v_{k-1}$. This proves Claim 1.]

[*Proof of Claim 2:* Let w be a linear combination of $v_1, v_2, \ldots, v_{k-1}$. We must show that w is a linear combination of $v_1, v_2, \ldots, v_k$.

We know that w is a linear combination of $v_1, v_2, \ldots, v_{k-1}$. In other words, there exist numbers $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ such that $w = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1}$. Thus,

$$\begin{aligned} w &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1} \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1} + 0 \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1} + 0 v_k. \end{aligned}$$

This shows that w is a linear combination of $v_1, v_2, \ldots, v_k$. This proves Claim 2.]

[*Remark:* In our proof of Claim 2, we did not use the assumption that $v_{k-1} = v_k$. More generally, it is always true that a linear combination of a subsequence of a given list of vectors must always be a linear combination of the whole sequence. It is the other direction that is usually not valid.] $\square$

Exercise 4. Let $n \in \mathbb{N}$. For each $k \in \{1, 2, \ldots, n\}$, we let e_k denote 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 $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.)

Explain why each vector in $\mathbb{R}^n$ is a linear combination of $e_1, e_2, \ldots, e_n$.

Solution to Exercise 4. Let v be any vector in $\mathbb{R}^n$. We must prove that v is a linear combination of $e_1, e_2, \ldots, e_n$.

Write v in the form $v = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$, where $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ are its coordinates. For each $i \in \{1, 2, \ldots, n\}$, we have

$$e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{(with the 1 being in the } i\text{-th position)}$$

and thus

$$\lambda_i e_i = \begin{pmatrix} \lambda_i 0 \\ \lambda_i 0 \\ \vdots \\ \lambda_i 0 \\ \lambda_i 1 \\ \lambda_i 0 \\ \vdots \\ \lambda_i 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(where the entry λ_i is in the i-th position). Hence,

$$\begin{aligned}
&\lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n \\
&= \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \lambda_n \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 + 0 + 0 + \cdots + 0 \\ 0 + \lambda_2 + 0 + \cdots + 0 \\ \vdots \\ 0 + 0 + 0 + \cdots + \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = v.
\end{aligned}$$

In other words, $v = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n$. Hence, v is a linear combination of $e_1, e_2, \ldots, e_n$. □

Definition 0.1 and Definition 0.2 make sense not only for vectors in $\mathbb{R}^n$. For example, we can replace "vectors in $\mathbb{R}^n$" by "polynomials" in both definitions, and obtain definitions for linear combinations and linear independence of polynomials (in one variable x, with real coefficients).[2] Thus:

- The polynomial $x^2 + 2x - 1$ is a linear combination of the polynomials x^2 and $(x-1)^2$, since
$$x^2 + 2x - 1 = 2 \cdot x^2 + (-1) \cdot (x-1)^2 .$$

- The polynomials x^2, $(x-1)^2$, $(x-2)^2$ and $(x-3)^2$ are linearly dependent, since there exists a nontrivial linear relation between them: If we take $\lambda_1 = 1$, $\lambda_2 = -3$, $\lambda_3 = 3$ and $\lambda_4 = -1$, then
$$\begin{aligned}
&\lambda_1 x^2 + \lambda_2 (x-1)^2 + \lambda_3 (x-2)^2 + \lambda_4 (x-3)^2 \\
&= 1 \cdot x^2 + (-3) \cdot (x-1)^2 + 3 \cdot (x-2)^2 + (-1) \cdot (x-3)^2 = 0.
\end{aligned}$$

- The polynomials x^2, $(x-1)^2$ and $(x-2)^2$ are linearly independent, because the only relation
$$\lambda_1 x^2 + \lambda_2 (x-1)^2 + \lambda_3 (x-2)^2 = 0 \qquad (6)$$
between them is the trivial one. How to be sure of this? One way is to assume that (6) holds, and try to derive $\lambda_1 = \lambda_2 = \lambda_3 = 0$ from it. This can be done, for example, by comparing coefficients: Straightforward expansion yields
$$\begin{aligned}
&\lambda_1 x^2 + \lambda_2 (x-1)^2 + \lambda_3 (x-2)^2 \\
&= \lambda_1 x^2 + \lambda_2 \left(x^2 - 2x + 1\right) + \lambda_3 \left(x^2 - 4x + 4\right) \\
&= (\lambda_1 + \lambda_2 + \lambda_3) x^2 + (-2\lambda_2 - 4\lambda_3) x + (\lambda_2 + 4\lambda_3),
\end{aligned}$$

[2] Of course, we then have to interpret 0 as the zero polynomial rather than the zero vector $0_{n\times 1}$.

and thus (6) rewrites as

$$(\lambda_1 + \lambda_2 + \lambda_3) x^2 + (-2\lambda_2 - 4\lambda_3) x + (\lambda_2 + 4\lambda_3) = 0.$$

Since equal polynomials have equal coefficients, this is equivalent to the following system of linear equations:

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ -2\lambda_2 - 4\lambda_3 = 0 \\ \lambda_2 + 4\lambda_3 = 0 \end{cases},$$

whose only solution is $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, 0)$ (you know how to check this) – and thus the relation must be trivial. So the polynomials x^2, $(x-1)^2$ and $(x-2)^2$ are linearly independent.

Exercise 5. **(a)** Are the four polynomials $x^3, (x-1)^3, (x-2)^3, (x-3)^3$ linearly independent?

(b) Is 1 a linear combination of these four polynomials?

Solution to Exercise 5. **(a)** Yes.

Proof. Let $ax^3 + b(x-1)^3 + c(x-2)^3 + d(x-3)^3 = 0$ be any relation between these four polynomials. We must show that this relation is trivial, i.e., that $a = b = c = d = 0$.

We have

$$\begin{aligned} & ax^3 + b \underbrace{(x-1)^3}_{=x^3-3x^2+3x-1} + c \underbrace{(x-2)^3}_{=x^3-6x^2+12x-8} + d \underbrace{(x-3)^3}_{=x^3-9x^2+27x-27} \\ &= ax^3 + b\left(x^3 - 3x^2 + 3x - 1\right) + c\left(x^3 - 6x^2 + 12x - 8\right) + d\left(x^3 - 9x^2 + 27x - 27\right) \\ &= ax^3 + bx^3 - 3bx^2 + 3bx - b + cx^3 - 6cx^2 + 12cx - 8c + dx^3 - 9dx^2 + 27dx - 27d \\ &= (a+b+c+d) x^3 + (-3b - 6c - 9d) x^2 + (3b + 12c + 27d) x + (-b - 8c - 27d). \end{aligned}$$

Thus, our relation $ax^3 + b(x-1)^3 + c(x-2)^3 + d(x-3)^3 = 0$ rewrites as the following equality:

$$(a+b+c+d) x^3 + (-3b - 6c - 9d) x^2 + (3b + 12c + 27d) x + (-b - 8c - 27d) = 0.$$

But a polynomial can only be 0 if all its coefficients are 0 (since a polynomial is defined as a sequence of coefficients). Thus, the equality that we have just showed entails that

$$\begin{aligned} a + b + c + d &= 0; \\ -3b - 6c - 9d &= 0; \\ 3b + 12c + 27d &= 0; \\ -b - 8c - 27d &= 0. \end{aligned}$$

This is a system of 4 linear equations in a, b, c, d. Solving it (e.g., by Gaussian elimination), we find that the only solution is $a = b = c = d = 0$. Thus, we conclude that $a = b = c = d = 0$, which is exactly what we needed to show.

[*Remark:* The claim generalizes: For any $n \in \mathbb{N}$ and any $n+1$ distinct reals $a_0, a_1, \ldots, a_n$, the $n+1$ polynomials $(x - a_0)^n, (x - a_1)^n, \ldots, (x - a_n)^n$ are linearly independent. To prove this, we would need some more advanced tools (such as the Vandermonde determinant), since the system of linear equations would get larger and larger as n grows.]

(b) Yes.

Proof. We have

$$\frac{1}{6}x^3 + \frac{-1}{2}(x-1)^3 + \frac{1}{2}(x-2)^3 + \frac{-1}{6}(x-3)^3 = 1. \tag{7}$$

[*Remark:* How did we find this relation? Well, we need to find four numbers a, b, c, d such that

$$ax^3 + b(x-1)^3 + c(x-2)^3 + d(x-3)^3 = 1.$$

We can expand the left hand side of this equation (as we already did in the proof of part **(a)**), thus rewriting the equation as

$$(a+b+c+d)x^3 + (-3b - 6c - 9d)x^2 + (3b + 12c + 27d)x + (-b - 8c - 81d) = 1.$$

But two polynomials are equal if and only if their respective coefficients are equal (since a polynomial is defined as a sequence of coefficients). Thus, our equation amounts to the following system of 4 linear equations in a, b, c, d:

$$\begin{aligned} a + b + c + d &= 0; \\ -3b - 6c - 9d &= 0; \\ 3b + 12c + 27d &= 0; \\ -b - 8c - 81d &= 1. \end{aligned}$$

Now, Gaussian elimination shows that its solution is $a = \frac{1}{6}$, $b = \frac{-1}{2}$, $c = \frac{1}{2}$ and $d = \frac{-1}{6}$. This leads exactly to (7).]

[*Another remark:* If you know the notion of factorials ($n!$), then you can generalize (7): For any nonnegative integer n, we have

$$\sum_{k=0}^{n} \frac{(-1)^k}{k! \cdot (n-k)!}(x-k)^n = 1.$$

Or, rewritten without the summation sign:

$$\frac{(-1)^0}{0! \cdot (n-0)!}(x-0)^n + \frac{(-1)^1}{1! \cdot (n-1)!}(x-1)^n + \frac{(-1)^2}{2! \cdot (n-2)!}(x-2)^n + \cdots + \frac{(-1)^n}{n! \cdot (n-n)!}(x-n)^n = 1.$$

Proving this for all n is a different matter, however; this is often done in courses on enumerative combinatorics[3].] □

Similarly, we can define linear combinations and linear independence of matrices: Just replace "vectors in $\mathbb{R}^n$" by "$n \times m$-matrices" in Definition 0.1 and Definition 0.2[4]. For example, the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is a linear combination of the matrices $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, since

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Exercise 6. Consider the four matrices

$$N = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad E = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \qquad W = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

(The letter E stands for "east", not for the identity matrix, which some also call E.)

(a) Is the identity matrix I_2 a linear combination of N, E, S, W ?

(b) Are N, E, S, W linearly dependent?

Solution to Exercise 6. **(a)** No.

Proof. Each of the four matrices N, E, S, W has the property that the sum of the two diagonal entries equals the sum of the two off-diagonal entries[5]. Thus, every linear combination of N, E, S, W has this property as well[6]. But I_2 does not have this property.

(b) Yes. Indeed, $N + S = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = E + W$; in other words, $1N + (-1)E + 1S + (-1)W = 0_{2\times 2}$. This is a nontrivial relation between N, E, S, W.

[*Remark:* These answers can be found without lucky inspiration: For part **(a)**, you want to write I_2 in the form

$$I_2 = aN + bE + cS + dW \qquad \text{for some } a, b, c, d \in \mathbb{R};$$

[3]Combinatorialists usually restate this equality in the form

$$(-1)^0 \binom{n}{0} (x-0)^n + (-1)^1 \binom{n}{1} (x-1)^n + \cdots + (-1)^n \binom{n}{n} (x-n)^n = n!.$$

It is not hard to derive it from the result in `https://math.stackexchange.com/a/1943039/` .

[4]Of course, we then have to interpret 0 as the zero matrix $0_{n\times m}$ rather than the zero vector $0_{n\times 1}$.

[5]For example, for E, this holds because $0 + 1 = 1 + 0$.

[6]Make sure you understand why!

this would entail

$$\begin{aligned}
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= I_2 = aN + bE + cS + dW \\
&= a\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + d\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} a\cdot 1 + b\cdot 0 + c\cdot 0 + d\cdot 1 & a\cdot 1 + b\cdot 1 + c\cdot 0 + d\cdot 0 \\ a\cdot 0 + b\cdot 0 + c\cdot 1 + d\cdot 1 & a\cdot 0 + b\cdot 1 + c\cdot 1 + d\cdot 0 \end{pmatrix} = \begin{pmatrix} a+d & a+b \\ c+d & b+c \end{pmatrix};
\end{aligned}$$

this would mean that the equations

$$1 = a + d, \qquad 0 = a + b, \qquad 0 = c + d, \qquad 1 = b + c$$

hold (because two matrices are equal if and only if their respective entries are equal); but this is a system of linear equations that has no solutions. For part **(b)**, you are similarly trying to solve the equation $0_{2\times 2} = aN + bE + cS + dW$, but this time you are looking for a nontrivial solution.] □

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/`