Chapter 1: Vector Spaces

Juan Patricio Carrizales Torres

Aug 17, 2022

Problem 1.1. Let $\mathcal{V}$ be a vector space over $\mathbb{F}$. Show that if $\alpha, \beta \in \mathbb{F}$ and if $\mathbf{v}$ is a nonzero vector in $\mathcal{V}$, then $\alpha\mathbf{v} = \beta\mathbf{v} \implies \alpha = \beta$. [HINT: $\alpha - \beta \neq 0 \implies \mathbf{v} = (\alpha-\beta)^{-1}(\alpha-\beta)\mathbf{v}$.]

Proof. Suppose, to the contrary, that there are distinct $\alpha, \beta \in \mathbb{F}$ such that for some nonzero $\mathbf{v} \in \mathcal{V}$ we have $\alpha\mathbf{v} = \beta\mathbf{v}$. Then, $\alpha - \beta \neq 0$ and so $\mathbf{v} = (\alpha-\beta)^{-1}(\alpha-\beta)\mathbf{v}$. Hence,

$$\mathbf{v} = (\alpha-\beta)^{-1}\alpha\mathbf{v} - (\alpha-\beta)^{-1}\beta\mathbf{v} = (\alpha-\beta)^{-1}(\alpha\mathbf{v} - \beta\mathbf{v}).$$

Since $\alpha\mathbf{v} = \beta\mathbf{v}$, it follows that $\alpha\mathbf{v} - \beta\mathbf{v} = \beta\mathbf{v} - \beta\mathbf{v} = \mathbf{0}$. This implies that $\mathbf{v} = (\alpha-\beta)^{-1}\mathbf{0} = \mathbf{0}$. This is a contradiction to our assumption that $\mathbf{v}$ was nonzero.

Another way to prove this directly is by using the fact, for some $\alpha \in \mathbb{F}$ and nonzero vector $\mathbf{v}$, that $\alpha\mathbf{v} = \mathbf{0} \implies \alpha = 0$. A proof reads as follows:

Let $\alpha, \beta \in \mathbb{F}$ and $\mathbf{v} \in \mathcal{V}$ be some nonzero vector such that $\alpha\mathbf{v} = \beta\mathbf{v}$. Then, $\alpha\mathbf{v} - \beta\mathbf{v} = \beta\mathbf{v} - \beta\mathbf{v} = \mathbf{0}$ and so $(\alpha - \beta)\mathbf{v} = \mathbf{0}$. Since $\mathbf{v}$ is nonzero, it follows that $\alpha - \beta = 0$ and so $\alpha = \beta$. $\square$

Problem 1.2. Show that the space $\mathbb{R}^3$ endowed with the rule

$$\mathbf{x} \,\square\, \mathbf{y} = \begin{bmatrix} \max(x_1, y_1) \\ \max(x_2, y_2) \\ \max(x_3, y_3) \end{bmatrix}$$

for vector addition and the usual rule for scalar multiplication is not a vector space over $\mathbb{R}$.

Proof. We show that this space has no unique additive identity. Consider some $\mathbf{x} = (x_1, x_2, x_3)$. Then, both $\mathbf{y} = (x_1 - 1, x_2 - 1, x_3 - 1)$ and $\mathbf{z} = (x_1 - 2, x_2 - 2, x_3 - 2)$ are in $\mathbb{R}^3$ and they are distinct. Note that $\mathbf{x} \,\square\, \mathbf{y} = \mathbf{x}$ and $\mathbf{x} \,\square\, \mathbf{z} = \mathbf{x}$.

In fact, one can easily show that there is no vector that is an additive inverse of every vector (the zero vector $\mathbf{0}$) since one can easily construct a vector with elements lower than the ones from any other vector. $\square$

Problem 1.3. Let $\mathcal{C} \subset \mathbb{R}^3$ denote the set of vectors $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ such that th polynomial $a_1 + a_2 t + a_3 t^2 \geq 0$ for every $t \in \mathbb{R}$. Show that it is closed under vector addition (i.e., $\mathbf{a}, \mathbf{b} \in \mathcal{C} \implies \mathbf{a} + \mathbf{b} \in \mathcal{C}$), but that $\mathcal{C}$ is not a vector space over $\mathbb{R}$. [REMARK: A set $\mathcal{C}$ with the indicated two properties is called a **cone**.]

Proof. We first show that $\mathcal{C}$ is closed under addition. Consider any $\mathbf{a}, \mathbf{b} \in \mathcal{C}$. Then, for every $t \in \mathbb{R}$ we have $a_1 + a_2 t + a_3 t^2 \geq 0$ and $b_1 + b_2 t + b_3 t^2 \geq 0$. Then,

$$a_1 + a_2 t + a_3 t^2 + b_1 + b_2 t + b_3 t^2 = (a_1 + b_1) + (a_2 + b_2)t + (a_3 + b_3)t^2 \geq 0$$

for every $t \in \mathbb{R}$. Thus, $\mathbf{a} + \mathbf{b} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{bmatrix} \in \mathcal{C}$. However, it is not close under scalar multiplication. Consider some nonzero $\mathbf{v} \in \mathcal{C}$ and let $\alpha = -1$. Since $v_1 + v_2 t + v_3 t^2 \geq 0$ for every $t \in \mathbb{R}$, it follows that $(-v_1) + (-v_2)t + (-v_3)t^2 < 0$ for every positive t. Hence, $(-1)\mathbf{v} \notin \mathcal{C}$ and so it is not a vector space over $\mathbb{R}$. $\square$

Problem 1.5. Let $\mathcal{F}$ denote the set of continuous real-valued functions $f(x)$ on the interval $0 \leq x \leq 1$. Show that $\mathcal{F}$ is a vector space over $\mathbb{R}$ with respect to the natural rules of vector addition $((f_1 + f_2)(x) = f_1(x) + f_2(x))$ and scalar multiplication $((\alpha f)(x) = \alpha f(x))$.

Proof.

(a) **Closed under vector addition**

Consider two functions $f, g \in \mathcal{F}$. Let $x \in [0,1]$. Then, $f(x), g(x) \in \mathbb{R}$ and so $(f + g)(x) = f(x) + g(x) \in \mathbb{R}$ since $\mathbb{R}$ is closed under addition. Therefore, $f + g$ is a real-valued function on the interval $[0,1]$ and so $(f + g) \in \mathcal{F}$.

(b) **Closed under scalar multiplication**

Consider some function $f \in \mathcal{F}$ and real number α. Let $x \in [0,1]$. Then, $f(x) \in \mathbb{R}$ and so $(\alpha f)(x) = \alpha f(x) \in \mathbb{R}$ since $\mathbb{R}$ is closed under multiplication. Thus, αf is a real-valued function on the interval $[0,1]$ and so $\alpha f \in \mathcal{F}$.

(c) **Vector addition is commutative**

Let $f, g \in \mathcal{F}$ and $x \in [0,1]$. Then, $(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)$ since addition in the set of real numbers is commutative.

(d) **Vector addition is associative**

Let $f, g, h \in \mathcal{F}$ and $x \in [0,1]$. Then, $((f+g)+h)(x) = (f+g)(x) + h(x) = f(x) + g(x) + h(x) = f(x) + (g+h)(x) = (f + (g+h))(x)$ since addition in $\mathbb{R}$ is associative (the order of addition does not matter).

(e) **Existence of additive identity**

Let $f : [0,1] \to \mathbb{R}$ be defined by $f(x) = 0$ for all $x \in [0,1]$. Then, f is a continous real-valued function and so $f \in \mathcal{F}$. Consider any $g \in \mathcal{F}$ and let $a \in [0,1]$. Then, $(f+g)(a) = f(a) + g(a) = 0 + g(a) = g(a)$ since 0 is the additive identity of real numbers. Thus, f is an additive identitive in $\mathcal{F}$.

(f) **Existence of additive inverse**

Consider some $f \in \mathcal{F}$. Let $g : [0,1] \to \mathbb{R}$ be defined by $g(x) = -f(x)$ for all $x \in [0,1]$. Consider some $x \in [0,1]$ and so $(f+g)(x) = f(x) + g(x) = f(x) - f(x) = 0$. Hence, g is the additive inverse of f.

(g) $f \in \mathcal{F} \implies (1)f = f$

Let $f \in \mathcal{F}$. Consider any $x \in [0,1]$ and so $f(x) = (1)f(x)$. Thus, $f = (1)f$.

(h) For any $\alpha, \beta \in \mathbb{R}$ and vector $f \in \mathcal{F}$, $\alpha(\beta f) = (\alpha\beta)f$

Let $f \in \mathcal{F}$ and $\alpha, \beta \in \mathbb{R}$. Consider any $x \in [0,1]$ and so $\alpha(\beta f)(x) = \alpha(\beta f(x)) = (\alpha\beta)f(x)$ since multiplication in $\mathbb{R}$ is associative. Thus, $\alpha(\beta f) = (\alpha\beta)f$

(i) For any $\alpha, \beta \in \mathbb{R}$ and vector $f \in \mathcal{F}$, $(\alpha+\beta)f = \alpha f + \beta f$

Let $f \in \mathcal{F}$ and $\alpha, \beta \in \mathbb{R}$. Consider any $x \in [0,1]$ and so $(\alpha+\beta)f(x) = \alpha f(x) + \beta f(x) = (\alpha f + \beta f)(x)$ since multiplication over addition is distributive for real numbers.

□

Lemma 1. Let $\mathcal{S}$ be a nonempty subset of a vector space $\mathcal{M}$ over $\mathbb{F}$. Then, $\mathcal{S}$ is a vector space if and only if for every pair of vectors $\mathbf{v}, \mathbf{a} \in \mathcal{S}$ and $\alpha, \beta \in \mathbb{F}$, $\alpha\mathbf{v} + \beta\mathbf{a} \in \mathcal{S}$.

Proof. Assume that $\mathcal{S}$ is a vector space and so it is closed under addition and scalar multiplication. Let $\mathbf{v}, \mathbf{a} \in \mathcal{S}$ and $\alpha, \beta \in \mathbb{F}$, then $\alpha\mathbf{v}, \beta\mathbf{a} \in \mathcal{S}$ and so $\alpha\mathbf{v} + \beta\mathbf{a} \in \mathcal{S}$.
Suppose, for every pair of vectors $\mathbf{v}, \mathbf{a} \in \mathcal{S}$ and $\alpha, \beta \in \mathbb{F}$, that $\alpha\mathbf{v} + \beta\mathbf{a} \in \mathcal{S}$. Let $\alpha = 0$ and $\beta \in \mathbb{F}$. Consider any vectors $\mathbf{v}, \mathbf{a} \in \mathcal{S}$. Then, $\alpha\mathbf{v} = 0$ is the additive identity of $\mathcal{M}$ and so $\beta\mathbf{a} = \alpha\mathbf{v} + \beta\mathbf{a} \in \mathcal{S}$. Thus, $\mathcal{S}$ is closed under scalar multiplication.
Consider some vectors $\mathbf{v}, \mathbf{a} \in \mathcal{S}$ and let $\alpha = \beta = 1$. Then, $\mathbf{v}+\mathbf{a} = (1)\mathbf{v} + (1)\mathbf{a} = \alpha\mathbf{v} + \beta\mathbf{a} \in \mathcal{S}$ since $\mathbf{v}, \mathbf{a} \in \mathcal{M}$. Therefore, $\mathcal{S}$ is closed under addition and so it is a vector space. □

Problem 1.6. Let F_0 denote the set of continuous real-valued functions $f(x)$ on the interval $0 \le x \le 1$ that met the auxiliary constraints $f(0) = 0$ and $f(1) = 0$. Show that F_0 is a vector space over $\mathbb{R}$ with respect to the natural rules of vector addition and scalar multiplication that were introducd in **Excercise 1.5** and that F_0 is a subspace of the vector space $\mathcal{F}$ that was considered there.

Proof. By definition, $F_0 \subseteq \mathcal{F}$. Let's prove that it is closd under addition and scalar multiplication. Consider some $f, g \in F_0$ and $\alpha, \beta \in \mathbb{R}$. Then, $\alpha f + \beta g$ is a real-valued function since $f, g \in \mathcal{F}$. Particularly, $(\alpha f + \beta g)(0) = \alpha f(0) + \beta g(0) = 0 + 0 = \alpha f(1) + \beta g(1) = (\alpha f + \beta g)(1)$ and so, by condition, it is a vector in F_0. Therefore, F_0 is a subspace of $\mathcal{F}$. □

Problem 1.7. Let F_1 denote the set of continuous real-valued functions $f(x)$ on the interval $0 \le x \le 1$ that meet the auxiliary constraints $f(0) = 0$ and $f(1) = 1$. Show that F_1 is not a vector space over $\mathbb{R}$ with respect to the natural rules of vector addition and scalar multiplication that were introduced in **Exercise 1.5**.

Proof. We know that $F_1 \subseteq \mathcal{F}$. Consider some $f \in F_1$. Then, $(2)f$ is a continuous real-valued function since $f \in \mathcal{F}$. However, note that $(2f)(1) = (2)f(1) = 2 \neq 1$ and so $(2)f \notin F_1$. Hence, F_1 is not closed under scalar multiplication and so F_1 is not a subspace of $\mathcal{F}$. □

Problem 1.8. Verify the last assertion; i.e., if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_k\}$ is a set of linearly independent vectors in the space $\mathcal{V}$ over $\mathbb{F}$ and if $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_k\mathbf{v}_k$, where $\alpha_j, \beta_j \in \mathbb{F}$ for $j = 1, \ldots, k$, then $\alpha_j = \beta_j$ for $j = 1, \ldots, k$.

Proof. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of linearly independet vectors in the space $\mathcal{V}$ over $\mathbb{F}$. Furthermore, assume that there is some vector $\mathbf{v} \in \mathcal{V}$ such that $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_k\mathbf{v}_k$, where $\alpha_j, \beta_j \in \mathbb{F}$ for $j = 1, \ldots, k$. Because $\mathbf{v} \in \mathcal{V}$, it follows that

$$\begin{aligned}
\mathbf{v} - \mathbf{v} &= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k - (\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_k\mathbf{v}_k) \\
&= (\alpha_1 - \beta_1)\mathbf{v}_1 + (\alpha_2 - \beta_2)\mathbf{v}_2 + \cdots + (\alpha_k - \beta_k)\mathbf{v}_k = 0.
\end{aligned}$$

Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent, it follows that $\alpha_j - \beta_j = 0$ and so $\alpha_j = \beta_j$ for $j = 1, \ldots, k$. $\square$

Problem 1.10. Show that if

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix}$$

then

$$AB = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \end{bmatrix} B + \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \end{bmatrix} B + \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \end{bmatrix} B$$

and hence that

$$AB = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \begin{bmatrix} b_{11} & \ldots & b_{14} \end{bmatrix} + \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \begin{bmatrix} b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \begin{bmatrix} b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix}$$

Proof. By the definition of addition of matrices

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \end{bmatrix}$$

and so

$$AB = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \end{bmatrix} B + \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \end{bmatrix} B + \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \end{bmatrix} B$$

since the multiplication of matrices is distributive over addition. By the definition of matrix multiplication, each entry $c_{kl} = \sum_{j=1}^{q} a_{kj}b_{jl}$, for the rows $k = 1, \ldots, p$ and columns $l = 1, \ldots, r$. Note that each matrix component of A has just one nonzero column m and so each entry $c_{kl} = a_{km}b_{ml}$. Thus

$$\begin{aligned}
AB &= \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{11}b_{14} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{21}b_{14} \end{bmatrix} + \begin{bmatrix} a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} & a_{12}b_{24} \\ a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} & a_{22}b_{24} \end{bmatrix} \\
&\quad + \begin{bmatrix} a_{13}b_{31} & a_{13}b_{32} & a_{13}b_{33} & a_{13}b_{34} \\ a_{23}b_{31} & a_{23}b_{32} & a_{23}b_{33} & a_{23}b_{34} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \begin{bmatrix} b_{11} & \ldots & b_{14} \end{bmatrix} + \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \begin{bmatrix} b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \begin{bmatrix} b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix}
\end{aligned}$$

$\square$

Problem 1.12. Show that if A and B are invertible matrices of the same size, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. Let $A, B \in \mathbb{F}^{p\times p}$ be invertible matrices. Then, A^{-1} and B^{-1} are left-right inverses of A and B, respectively. Therefore,

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1}\\ &= A(I_pA^{-1}) = AA^{-1}\\ &= I_p\end{aligned}$$

and

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B\\ &= B^{-1}(I_pB) = B^{-1}B\\ &= I_p,\end{aligned}$$

since matrix multiplication is associative. Thus, $B^{-1}A^{-1}$ is the **inverse** of AB and so AB is invertible. □

Problem 1.13. Show that the matrix $A = \begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 1 & 1 & 0\end{bmatrix}$ has no left inverses and no right inverses.

Proof. Suppose to the contrary, that A has some right inverse B. Then, $B \in \mathbb{F}^{3\times 3}$ and $AB = C = I_3$. Therefore, $c_{22} = 1\cdot b_{12} + 1\cdot b_{22} + 0\cdot b_{32} = 1$ and $c_{32} = 1\cdot b_{12} + 1\cdot b_{22} + 0\cdot b_{32} = 0$, which is a contradiction.
Now, assume, to the contrary, that A has some left inverse B. Then, $B \in \mathbb{F}^{3\times 3}$ and $BA = C = I_3$. Hence, $c_{11} = b_{11}\cdot 1 + b_{12}\cdot 1 + b_{13}\cdot 1 = 1$, $c_{12} = b_{11}\cdot 0 + b_{12}\cdot 1 + b_{13}\cdot 1 = 0$ and $c_{13} = b_{11}\cdot 1 + b_{12}\cdot 0 + b_{13}\cdot 0 = 0$. This leads to the contradiction $1 = 0$. □

Problem 1.15. Show that if a matrix $A \in \mathbb{C}^{p\times q}$ has two right inverse B_1 and B_2, then $\lambda B_1 + (1-\lambda)B_2$ is also a right inverse for every choice of $\lambda \in \mathbb{C}$.

Proof. Suppose that A has two right inverses $B_1, B_2 \in \mathbb{C}^{q\times p}$. Choose any $\lambda \in \mathbb{C}$. Then

$$\begin{aligned}A\left(\lambda B_1 + (1-\lambda)B_2\right) &= \lambda AB_1 + (1-\lambda)AB_2\\ &= \lambda I_p + (1-\lambda)I_q = (\lambda-\lambda)I_p + I_p\\ &= I_p.\end{aligned}$$

since matrix multiplication is distributive and under scalar multiplication is commutative. Assuming that another matrix A' has two left inverses $B_1, B_2 \in \mathbb{C}^{q\times p}$ and let $\lambda \in \mathbb{C}$. Then,

$$\begin{aligned}\left(\lambda B_1 + (1-\lambda)B_2\right)A &= \lambda B_1A + (1-\lambda)B_2A\\ &= \lambda I_q + (1-\lambda)I_q = (\lambda-\lambda)I_q + I_q\\ &= I_q.\end{aligned}$$

□

Problem 1.16. Show that a given matrix $A \in \mathbb{F}^{p\times q}$ has either 0, 1 or infinitely many right inverses and that the same conclusion prevails for left inverses.

Proof. Consider the vector space $\mathbb{F}^{p\times q}$ with $p, q \geq 2$. Consider the zero matrix $\mathbf{0} \in \mathbb{F}^{p\times q}$ and so it has no left and right invertibles since $A\mathbf{0} = \mathbf{0}B = \mathbf{0}$ for all $A, B \in \mathbb{F}^{q\times p}$.
Now, let's construct some matrix $A \in \mathbf{F}^{p\times q}$. Now, let each entry $a_{ii} = 1$ while the other be zero. For instance, in the case $p > q$, we have that

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

If $p \geq q$ (greater or equal number of rows), then $A^T A = I_p$. On the other hand, if $q \geq p$ (greater or equal numbr of columns), then $AA^T = I_q$.

Hence, any matrix $A \in \mathbb{F}^{p\times q}$ can have 0 or at least one right/left invertible (depending on the order relation of rows and columns). If it has more than one right/left invertibles, then one can construct and infinity of right/left invertibles with the formula given in **Problem 1.15**.

$\square$

Problem 1.19. Show that if T is a linear transformation from a vector space $\mathcal{U}$ over $\mathbb{F}$ with basis $\{\mathbf{u}_1, \dots, \mathbf{u}_q\}$ into a vector space $\mathcal{V}$ over $\mathbb{F}$ with basis $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$, then there exists a unique st of scalars $a_{ij} \in \mathbb{F}$, $i = 1, \dots, p$ and $j = 1, \dots, q$ such that

$$T\mathbf{u}_j = \sum_{i=1}^{p} a_{ij}\mathbf{v}_i \ \text{ for } \ j = 1, \dots, q \tag{1}$$

and hence that

$$T\left(\sum_{j=1}^{q} x_j \mathbf{u}_j\right) = \sum_{i=1}^{p} y_i \mathbf{v}_i \iff A\mathbf{x} = \mathbf{y},$$

where $\mathbf{x} \in \mathbb{F}^q$ has components $x_1, \dots, x_q$, $\mathbf{y} \in \mathbb{F}^p$ has components $y_1, \dots, y_p$ and the entries a_{ij} of $A \in \mathbb{F}^{p\times q}$ are determined by formula 1.

Proof. Since T is a linear transformation, T maps u_j ($j \in \{1, \dots, q\}$ into only one vector $b \in \mathcal{V}$. Because $\mathcal{V}$ has basis $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and $T\mathbf{u}_j \in \mathcal{V}$ represents only one vector, it follows that there is a unique set of scalars $a_{ij} \in \mathbb{F}$ such that

$$T\mathbf{u}_j = \sum_{i=1}^{p} a_{ij}\mathbf{v}_i.$$

Note that it is possible for $T\mathbf{u}_i = T\mathbf{u}_j$, where $i \neq j$ (Non-injective linear transformation). However, each are still represented by a unique set of scalars. Now, assume that

$$T\left(\sum_{j=1}^{q} x_j \mathbf{u}_j\right) = \sum_{i=1}^{p} y_i \mathbf{v}_i.$$

Note that $\sum_{j=1}^{q} x_j \mathbf{u}_j \in \mathcal{U}$ and $\sum_{i=1}^{p} y_i \mathbf{v}_i \in \mathcal{V}$, since they are linear combinations of the basis of $\mathcal{U}$ and $\mathcal{V}$, respectively. Furthermore,

$$\begin{aligned} T\left(\sum_{j=1}^{q} x_j \mathbf{u}_j\right) &= \sum_{j=1}^{q} x_j T(\mathbf{u}_j) \\ &= \sum_{j=1}^{q} x_j \left(\sum_{i=1}^{p} a_{ij} \mathbf{v}_i\right) \\ &= \sum_{i=1}^{p} \sum_{j=1}^{q} x_j a_{ij} \mathbf{v}_i = \sum_{i=1}^{p} y_i \mathbf{v}_i \end{aligned}$$

due to the linearity of T (Note that both characteristics of linear mappings $T(\mathbf{a}+\mathbf{b}) = T\mathbf{a} + T\mathbf{b}$ and $T(\alpha\mathbf{a}) = \alpha T\mathbf{a}$ are used). If we fix i, then each basis vector is expressed with their coefficient as

$$\sum_{j=1}^{q} x_j a_{ij} \mathbf{v}_i = \beta_i \mathbf{v}_i \qquad \text{and} \qquad y_i \mathbf{v}_i = \alpha_i \mathbf{v}_i.$$

Recall that $\sum_{j=1}^{p} \beta_i \mathbf{v}_i = \sum_{j}^{p} \alpha_i \mathbf{v}_i$. Hence, $\beta_i = \alpha_i$ for all $i = 1, \ldots, p$. Now consider the $p \times q$ matrix

$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \ldots & \vec{a}_q \end{bmatrix},$$

where $\vec{a}_j = (a_{1j}, a_{2j}, \ldots, a_{pj})^T$ is a column vector. Now, let $\mathbf{x} = (x_1, \ldots, x_q)^T$ be a column matrix. Then,

$$A\mathbf{x} = \sum_{j=1}^{q} \vec{a}_j x_j = \begin{bmatrix} \sum_{j=1}^{q} x_j a_{1j} \\ \sum_{j=1}^{q} x_j a_{2j} \\ \vdots \\ \sum_{j=1}^{q} x_j a_{pj} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \mathbf{y},$$

assuming each $\mathbf{v}_i$ has one as the only nonzero entry in the i'th row. For the converse, assume that

$$A\mathbf{x} = \sum_{j=1}^{q} \vec{a}_j x_j = \begin{bmatrix} \sum_{j=1}^{q} x_j a_{1j} \\ \sum_{j=1}^{q} x_j a_{2j} \\ \vdots \\ \sum_{j=1}^{q} x_j a_{pj} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \mathbf{y}.$$

Then, $\sum_{j=1}^{q} x_j a_{ij} = y_i$. This implies that

$$\begin{aligned}
\sum_{i=1}^{p}\sum_{j=1}^{q} x_j a_{ij}\mathbf{v}_i &= \sum_{j=1}^{p} x_j \sum_{i=1}^{p} a_{ij}\mathbf{v}_i \\
&= \sum_{j=1}^{p} x_j T(\mathbf{u}_j) = \sum_{j=1}^{p} T(x_j\mathbf{u}_j) \\
&= T\left(\sum_{j=1}^{p} x_j\mathbf{u}_j\right) = \sum_{i=1}^{p} y_i\mathbf{v}_i.
\end{aligned}$$

□

Problem 1.21. Show that if $A \in \mathbb{C}^{n\times n}$ and $A^k = O_{n\times n}$ for some positive integer k, then $I_n - A$ is invertible. [The author gaves use the inverse B, we just showed that it was an inverse].

Proof. Let $B = I_n + A + A^2 + \cdots + A^{k-2} + A^{k-1}$. Then,

$$\begin{aligned}
B(I_n - A) &= BI_n - BA \\
&= \left(I_n + A + A^2 + \cdots + A^{k-2} + A^{k-1}\right) - \left(A + A^2 + \cdots + A^{k-2} + A^{k-1} + A^k\right) \\
&= (I_n - A)B = I_n - A^k = I_n
\end{aligned}$$

since $A^k = O_{n\times n}$. Hence, $I_n - A$ is invertible and B is its inverse. □

Problem 1.22. Show that even though all the diagonal entries of the matrix

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

are equal to zero, A is invertible, and find A^{-1}.

Proof. Let

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

It is readily checked that

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

Hence, A is invertible and $A^{-1} = B$, despite the fact that all diagonal entries of A are 0. □

Problem 1.23. Use Excercise 1.21 to show that a triangular $n \times n$ matrix A with nonzero diagonal entries is invertible by writing

$$A = D + (A - D) = D(I_n + D^{-1}(A - D)),$$

where D is the diagonal matrix with $d_{jj} = a_{jj}$ for $j = 1, \dots, n$.

Proof. Let $A \in \mathbb{F}^{n \times n}$ be an upper triangular matrix. Note that $D^{-1}(D - A) = -D^{-1}A + I_n$ and $-D^{-1}A$ is an upper triangular matrix with a diagonal of ones. Thus, $C_1 = -D^{-1}A + I_n$ is a diagonix of level 1 and so $C_1^n = C_{n \cdot 1} = O_{n \times n}$ (**Lemma Diagonix**). Then, $M = I_n + D^{-1}(A - D) = I_n - D^{-1}(D - A)$ is an invertible matrix by **Problem 1.21**. Therefore, $A = D(I_n + D^{-1}(A - D)) = DM$ is invertible and its inverse is $M^{-1}D^{-1}$. For the case of lower triangular matrices, we just take the transpose of our upper one, namely,

$$A^T = (I_n + D^{-1}(A - D))^T D^T = M^T D.$$

The lower triangular matrix $A^T \in \mathbb{F}^{n \times n}$ with nonzero diagonal entries is invertible and its inverse is $D^{-1}\left(M^{-1}\right)^T$. □

Problem 1.24. Let $A = \begin{bmatrix} A_{11} & A_{12} \\ O_{q \times p} & A_{22} \end{bmatrix}$ be an upper block triangular matrix with blocks A_{11} of size $p \times p$ and A_{22} of size $q \times q$. Show that A is invertible if and only if A_{11} and A_{22} are both invertible and that in this case

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O_{q \times p} & A_{22}^{-1} \end{bmatrix}.$$

Proof. First, assume that A_{11} and A_{22} are both invertible. Hence, the matrix $B = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O_{q \times p} & A_{22}^{-1} \end{bmatrix}$ is defined. Then,

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ O_{q \times p} & A_{22} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O_{q \times p} & A_{22}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}A_{11}^{-1} & -A_{12}A_{22}^{-1} + A_{12}A_{22}^{-1} \\ O_{q \times p} & A_{22}A_{22}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}A_{11}^{-1} & O_{p \times q} \\ O_{q \times p} & A_{22}A_{22}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} I_p & O_{p \times q} \\ O_{q \times p} & I_q \end{bmatrix} = I_{p+q} \end{aligned}$$

and

$$\begin{aligned} BA &= \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O_{q \times p} & A_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ O_{q \times p} & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^{-1}A_{11} & O_{p \times p} \\ O_{q \times q} & A_{22}^{-1}A_{22} \end{bmatrix} = \begin{bmatrix} I_p & O_{p \times p} \\ O_{q \times q} & I_q \end{bmatrix} = I_{p+q}. \end{aligned}$$

Hence, $B = A^{-1}$ and A is invertible.

For the converse, assume that A is invertible. Hence, there is some inverse $A^{-1} = B$. Hence,

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ O_{q\times p} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\ &= I_{p+q} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ O_{q\times p} & A_{22} \end{bmatrix} = BA \end{aligned}$$

Hence,

(a) $A_{11}B_{11} + A_{12}B_{21} = I_p$.

(b) $A_{11}B_{12} + A_{12}B_{22} = O_{p\times q}$.

(c) $O_{q\times p}B_{11} + A_{22}B_{21} = O_{q\times p}$.

(d) $O_{q\times p}B_{12} + A_{22}B_{22} = I_q$.

and

(a) $B_{11}A_{11} + B_{12}O_{q\times p} = I_p$.

(b) $B_{11}A_{12} + B_{12}A_{22} = O_{p\times q}$.

(c) $B_{21}A_{11} + B_{22}O_{q\times p} = O_{q\times p}$.

(d) $B_{21}A_{12} + B_{22}A_{22} = I_q$.

Thus, $A_{22}B_{21} = B_{21}A_{11} = O_{q\times p}$. Then,

$$(B_{11}A_{11})A_{11} = (I_p)A_{11} = (A_{11}B_{11} + A_{12}B_{21})A_{11}$$

□

1 INTERESTING LEMMAS

Lemma 1. Let $A \in \mathbb{F}^{p\times q}$. A is right-invertible if and only if the rows are linearly independent. The same can be said for left-invertibility and columns.

Proof. Assume that the rows of A are linearly independent. We show that we can construct a right-inverse $B \in \mathbb{F}^{q\times p}$. □

Lemma 2. Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ with basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Also, let $T : \mathcal{V} \to \mathcal{U}$ be a linear transformation, where $\mathcal{U}$ is a vector space over $\mathbb{F}$. Then, $T\mathbf{v}_j \neq T\mathbf{v}_i$ for all $i \neq j$.

Proof. Suppose, to the contrary, that there are two distinct basis vectors such that $T\mathbf{v}_i = T\mathbf{v}_j$. Then, $T\mathbf{v}_i - T\mathbf{v}_j = 0$ since they belong to a vector space. Because T is a linear transformation, it follows that T $\square$

Lemma Diagonix. A square matrix $A \in \mathbb{C}^{n\times n}$ is **upper diagonix** of level $\mathbf{k} \in \{1, 2, \ldots, n-1\}$ if $a_{ij} = 0$ when $j < i + k$. It can be seen as some type of upper triangular matrix whose diagonal extends further. For instance,

$$\begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

are upper diagonix matrices of levels 1, 3 and 1, respectively. We now proceed to give a proposition. Let $A, B \in \mathbb{F}^{n\times n}$ be an upper diagonix matrix of levels k_1 and k_2, respectively. They can be expressed as A_{k1} and B_{k2}. Then,

$$A_{k_1} B_{k_2} = C_{k_1+k_2},$$

namely, it equals an upper diagonix $C \in \mathbb{F}^{n\times n}$ of level $k_1 + k_2$. Note that if $n - 1 < k_1 + k_2$, then $C_{k_1+k_2} = O_{n\times n}$.

Proof. Since A_{k_1} and B_{k_2} are square matrices, it follows that $A_{k_1} B_{k_2} = C$ for some $C \in \mathbb{F}^{n\times n}$. We know that

$$c_{lm} = \sum_{i=1}^{n} a_{li} b_{im}.$$

If $m < l + (k_1 + k_2)$, then we have two cases.
Case1. If $i < l + k_1$, then $a_{li} = 0$.
Case2. If $i \geq l + k_1$, then $m < l + (k_1 + k_2) \leq i + k_2$ and so $b_{im} = 0$.
Thus, $b_{lm} = 0$ when $m \leq l + (k_1 + k_2)$. Hence, C is an upper diagonix of level $k_1 + k_2$. $\square$

Corollary Diagonix 1. Consider some sequence of $n \geq 2$ upper diagonix matrices $X_1, X_2, \ldots, X_n \in \mathbb{F}^{m\times m}$ of levels $M_1, M_2, \ldots, M_n$, respectively. Then,

$$\prod_{i=1}^{n} X_i = Y_{\sum_{i=1}^{n} M_i},$$

namely, equal to an upper diagonix $Y \in \mathbb{F}^{m\times m}$ of level $\sum_i^n M_i$.

Proof. By the previous lemma, we know that $X_{k_1} Y_{k_2} = Z_{k_1+k_2}$. Thus, the result is true for $n = 2$. Now, consider some $k \geq 2$ and assume for some sequence of upper diagonix matrices $A_1, A_2, \ldots, A_k \in \mathbb{F}^{m\times m}$ with respective levels $M_1, M_2, \ldots, M_k$ that

$$\prod_{i=1}^{k} A_i = C_{\sum_{i=1}^{k} M_i},$$

where $C \in \mathbb{F}^{m \times m}$. We show for some sequence of upper diagonix matrices $B_1, B_2, \ldots, B_{k+1}$ with respective levels $M_1, M_2, \ldots, M_{k+1}$ that

$$\prod_{i=1}^{k+1} B_i = D_{\sum_{i=1}^{k+1} M_i}.$$

Note that,

$$\begin{aligned}
\prod_{i=1}^{k+1} B_i &= \left(\prod_{i=1}^{k} B_i\right) B_{k+1} \\
&= \left(E_{\sum_{i=1}^{k} M_i}\right) B_{k+1} \\
&= D_{\sum_{i=1}^{k} M_i + M_{k+1}} = D_{\sum_{i=1}^{k+1} M_i}
\end{aligned}$$

due to the inductive hypothesis and the result proven in the previous lemma. Also, the matrices $E, D \in \mathbb{F}^{m \times m}$. By the Principle of Mathematical Induction, this result is true for $n \geq 2$. □