Chapter 1

Basic Notions

1. Vector Spaces

Exercise 1.1 Let $\mathbf{x} = (1,2,3)^{\mathrm{T}}$, $\mathbf{y} = (y_1, y_2, y_3)^{\mathrm{T}}$ and $\mathbf{z} = (4,2,1)^{\mathrm{T}}$. Compute $2\mathbf{x}$, $3\mathbf{y}$, $\mathbf{x} + 2\mathbf{y} - 3\mathbf{z}$.

Proof. Little calculation reveals that

$$2\mathbf{x} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \quad 3\mathbf{y} = \begin{pmatrix} 3y_1 \\ 3y_2 \\ 3y_3 \end{pmatrix}, \quad \mathbf{x} + 2\mathbf{y} - 3\mathbf{z} = \begin{pmatrix} 2y_1 - 11 \\ 2y_2 - 4 \\ 2y_3 \end{pmatrix}$$

$\square$

Exercise 1.2 Which of the following sets (with natural addition and multiplication by a scalar) are vector spaces? Justify your answers.

a). The set of all continuous functions on the interval $[0,1]$;

b). The set of all non-negative functions on the interval $[0,1]$;

c). The set of all polynomials of degree *exactly* n;

d). The set of all symmetric $n \times n$ matrices, i.e. the set of matrices $A = \{a_{j,k}\}_{j,k=1}^{n}$ such that $A^{\mathrm{T}} = A$.

Proof.

a). Let $\mathcal{C}[0,1]$ be the set of all continuous functions on $[0,1]$. For any $f, g \in \mathcal{C}[0,1]$ and $\alpha \in \mathbb{R}$, we define

$$(f+g)(x) := f(x) + g(x) \quad \text{and} \quad (\alpha f)(x) := \alpha \cdot f(x)$$

for each $x \in [0,1]$. Therefore, $(\mathcal{C}[0,1], +, \cdot)$ is closed under addition and scalar multiplication. It's immediate to see that all the eight properties of vector space are all satisfied, that is

- $f + g = g + f$
- $f + (g + h) = (f + g) + h$
- $f + 0 = f$
- $f + (-f) = 0$
- $1f = f$
- $\alpha(\beta f) = (\alpha\beta)f$
- $(\alpha + \beta)f = \alpha f + \beta f$
- $\alpha(f + g) = \alpha f + \beta g$

Note that the function $0 \in \mathcal{C}[0,1]$ such that $0(x) = 0$ for each $x \in [0,1]$.

b). Let $\mathcal{B}$ is the set of all non-negative functions on $[0,1]$. Then $(\mathcal{B}, +\cdot)$ is not a vector space because it's not closed under scalar multiplication, i.e. if $f \in \mathcal{B}$, hence $f > 0$ yet

$$-f = (-1) \cdot f < 0.$$

(Even if we restrict the scalar be to positive real numbers, this set still won't be a vector space, because it fails to have an inverse.)

c). Let $\mathcal{P}$ be the set of all polynomials of degree exactly n, then $(\mathcal{P}, +, \cdot)$ is *not* a vector space, because the addtive indentity is the polynomial 0. However, $0 \notin \mathcal{P}$.

d). Let $\text{sym}(n)$ be the set of all symmetric $n \times n$ matrices. The addition and scalar multiplication are defined as an *entrywise* operations. Hence, $\text{sym}(n)$ is closed under $(+)$ and $(\cdot)$. The additive indentity is the matrix

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \text{sym}(n).$$

We could easily see that all the eight properties of vector spaces are all satisfied.

$\square$

Exercise 1.3 True or false:

a). Every vector space contains a zero vector; (**True.**)

b). A vector space can have more than one zero vector; (**False.** The zero vector is unique.)

c). An $m \times n$ matrix has m rows and n columns; (**True.**)

d). If f and g are polynomials of degree n, then $f + g$ is also a polynomial of degree n. (**False.** consider t^n and $t - t^n$.)

e). If f and g are polynomials of degree atmost n, the $f + g$ is also a polynomial of degree atmost n. (**True.**)

Exercise 1.4 Prove that a zero vector $\mathbf{0}$ of a vector space V is unique.

Proof. Suppose that $\mathbf{a}$ and $\mathbf{b}$ are the zero vectors of V. From the *Axioms of Vector Space*, we obtain that

$$\begin{aligned} \mathbf{a} &= \mathbf{a}+\mathbf{b} && (\mathbf{b} \text{ is the zero vector}) \\ &= \mathbf{b}+\mathbf{a} && (\text{commutitativity}) \\ &= \mathbf{b} && (\mathbf{a} \text{ is the zero vector}) \end{aligned}$$

Hence, a zero vector of a vector space is unique, and we usually denote it by $\mathbf{0}$. □

Exercise 1.5 What is the zero matrix of the space $M_{2\times 3}$?

Answer. In the space $M_{2\times 3}$, the zero matrix is

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

□

Exercise 1.6 Prove that the additive inverse of a vector space is unique.

Proof. Let $\mathbf{a}$ be an arbitrary vector. Assume the $\mathbf{a}$ has two inverses, namely $\mathbf{x}$ and $\mathbf{y}$. Hence

$$\begin{aligned} \mathbf{x} &= \mathbf{x}+\mathbf{0} \\ &= \mathbf{x}+(\mathbf{a}+\mathbf{y}) && (\mathbf{y} \text{ is an inverse}) \\ &= (\mathbf{x}+\mathbf{a})+\mathbf{y} && (\text{associativity}) \\ &= \mathbf{0}+\mathbf{y} && (\mathbf{x} \text{ is an inverse}) \\ &= \mathbf{y}. \end{aligned}$$

Therefore, the inverse of any vector $\mathbf{a} \in V$ is unique, and is usually denoted by $-\mathbf{a}$. □

Exercise 1.7 Prove that $0\mathbf{v} = \mathbf{0}$ for any vector $\mathbf{v} \in V$.

Proof. Let $\mathbf{v} \in V$ and $\mathbf{b}$ is an inverse of $0\mathbf{v}$. Therefore,

$$\begin{aligned} 0 &= 0\mathbf{v}+b \\ &= (0+0)\mathbf{v}+b \\ &= (0\mathbf{v}+0\mathbf{v})+b && (\text{distributivity}) \\ &= 0\mathbf{v}+(0\mathbf{v}+b) && (\text{associativity}) \\ &= 0\mathbf{v}+\mathbf{0} && (\mathbf{b} \text{ is an inverse of } 0\mathbf{v}) \\ &= 0\mathbf{v} \end{aligned}$$

for any $\mathbf{v} \in V$. □

Exercise 1.8 Prove that for any vector $\mathbf{v}$ its additive inverse $-\mathbf{v}$ is given by $(-1)\mathbf{v}$.

Proof. As proved in the above exercise for any $\mathbf{v} \in V$,

$$\mathbf{0} = 0\mathbf{v} = (1-1)\mathbf{v} = \mathbf{v} + (-1)\mathbf{v}$$

where the last equallity derives from the distributive property. Because $-\mathbf{v}$ is the inverse of $\mathbf{v}$, then

$$\begin{aligned} -\mathbf{v} &= -\mathbf{v} + \mathbf{0} \\ &= -\mathbf{v} + \left[\mathbf{v} + (-1)\mathbf{v}\right] \\ &= (\underbrace{-\mathbf{v} + \mathbf{v}}_{\mathbf{0}}) + (-1)\mathbf{v} \\ &= (-1)\mathbf{v} \end{aligned}$$

as desired. ☐

2. Linear Combination, bases

Exercise 2.1 Find the basis in the space of 3×2 matrices $M_{3\times 2}$.

Answer. Consider the vectors:

$$\mathbf{e_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{e_2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{e_3} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{e_4} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{e_5} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{e_6} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and we're going to prove that the system of thses vectors are a basis. Any matrix

$$\mathbf{v} = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} \in M_{3\times 2}$$

can be represented as the combination $\mathbf{v} = a\mathbf{e_1} + b\mathbf{e_2} + c\mathbf{e_3} + d\mathbf{e_4} + e\mathbf{e_5} + f\mathbf{e_6}$ thus this system is generating. Next we're going to prove the uniqueness.

Suppose that there are $\hat{a}, \hat{b}, \ldots, \hat{f}$ with

$$\begin{aligned} &\mathbf{v} = \hat{a}\mathbf{e_1} + \hat{b}\mathbf{e_2} + \hat{c}\mathbf{e_3} + \hat{d}\mathbf{e_4} + \hat{e}\mathbf{e_5} + \hat{f}\mathbf{e_6} \\ \implies &\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} \hat{a} & \hat{b} \\ \hat{c} & \hat{d} \\ \hat{e} & \hat{f} \end{bmatrix} \end{aligned}$$

This implies that each corresponding entry is equals. Hence the representation is unique. Therefore this system is a basis.

☐

Exercise 2.2 True or false:

a). Any set containing a zero vector is linearly dependent;

b). A basis must contain $\mathbf{0}$;

c). subsets of linearly dependent sets are linearly dependent;

d). subsets of linearly independent sets are linearly independent;

e). if $\alpha_1 \mathbf{v_1} + \alpha_2 \mathbf{v_2} + \cdots + \alpha_n \mathbf{v_n} = 0$ then all scalars α_k are zero.

Answer.

a). **True.** because $\mathbf{0}$ can be represented as a linear combination of the other vectors (simply put all the scalars to 0).

b). **No.** if so, they must be linearly dependent, which is not a base.

c). **No.** Take for example the system of linearly dependent $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\}$ where $\mathbf{e_1} = (1,0)$, $\mathbf{e_2} = (0,1)$ and $\mathbf{e_3} = (1,1)$. The subset $\{\mathbf{e_1}, \mathbf{e_2}\}$ is a basis, which is clearly not linearly dependent.

d). **True.** Supppose that the system $\{\mathbf{v_1}, \dots, \mathbf{v_p}\}$ is a subset of the linearly independent system $\{\mathbf{v_1}, \dots, \mathbf{v_p}, \dots, \mathbf{v_n}\}$. Let α_k the real numbers such that $\alpha_1 \mathbf{v_1} + \cdots + \alpha_p \mathbf{v_p} = \mathbf{0}$ hence

$$\alpha_1 \mathbf{v_1} + \cdots + \alpha_p \mathbf{v_p} + 0\mathbf{v_{p+1}} + \cdots + 0\mathbf{v_n} = \mathbf{0}.$$

Because the system $\{\mathbf{v_1}, \dots, \mathbf{v_p}, \dots, \mathbf{v_n}\}$ is linearly independent, therefore all the scalars $\alpha_k = 0$. Thus, the system $\{\mathbf{v_1}, \dots, \mathbf{v_p}\}$ is also linearly independent.

e). **No.** Take, $\mathbf{e_1} = (2,2)$ and $\mathbf{e_2} = (1,1)$ for instance. We have $\mathbf{e_1} - 2\mathbf{e_2} = \mathbf{0}$ yet the scalars are non-zero.

□

Exercise 2.3 Recall, that a matrix is called *symmetric* if $A^T = A$. Write down a basis in the space of *symmetric* 2×2 matrices (there are many possible answers). How many elements are there in the basis.

Answer. We are going to prove that the system $\{\mathbf{d_1}, \mathbf{d_2}, \mathbf{e_1}\}$ where

$$\mathbf{d_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{d_2} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{e_1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

is a basis. Observe that any symmetric matrix

$$\mathbf{v} = \begin{bmatrix} d_1 & e_1 \\ e_1 & d_2 \end{bmatrix}$$

can be represented as $\mathbf{v} = d_1\mathbf{d_1} + d_2\mathbf{d_2} + e_1\mathbf{e_1}$, hence it's generating. Note that the equation

$$d_1\mathbf{d_1} + d_2\mathbf{d_2} + e_1\mathbf{e_1} = \mathbf{0}$$
$$\begin{bmatrix} d_1 & e_1 \\ e_1 & d_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

holds only when all the scalars are all zero. Hence the system is linearly independent. Thus, it's a basis.

□

Exercise 2.4 Write down a basis for the space of

a). 3×3 symmetric matrices;

b). $n \times n$ symmetric matrices;

c). $n \times n$ antisymmetric matrices.

Answer.

a). we are going to prove that the system of vectors

$$\mathbf{d_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{d_2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{d_3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{e_1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{e_2} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{e_3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

is the basis. First of, any symmetric matrix

$$\mathbf{v} = \begin{bmatrix} d_1 & e_1 & e_2 \\ e_1 & d_2 & e_3 \\ e_2 & e_3 & d_3 \end{bmatrix}$$

can be represented as

$$\mathbf{v} = d_1\mathbf{d_1} + d_2\mathbf{d_2} + d_3\mathbf{d_3} + e_1\mathbf{e_1} + e_2\mathbf{e_2} + e_3\mathbf{e_3}$$

yeilds that the system is generating. Similar to the previous problem, if the linear combination of these vectors equals $\mathbf{0}$, then all the scalars must equals zero. Thus it's linearly independent. Therefore it's a basis.

b). Working on it.

c). Working on it.

□

Exercise 2.5 Let a system of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ be linearly independent but not generating. Show that it is possible to find a vector $\mathbf{v}_{r+1}$ such that the system $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}$ is linearly independent.

Proof. Because the system $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ is not generating, therefore there exists a vector $\mathbf{v}_{r+1}$ such that $\mathbf{v}_{r+1}$ cannot be represented as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$. Let α_i be the scalars such that

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_r \mathbf{v}_r + \alpha_{r+1} \mathbf{v}_{r+1} = \mathbf{0} \tag{1.1}$$

Now we have to prove that all the scalars are all zero. If $\alpha_{r+1} \neq 0$ then

$$\mathbf{v}_{r+1} = -\sum_{i=1}^{r} \frac{\alpha_i}{\alpha_{r+1}} \cdot \mathbf{v}_i,$$

meaning $\mathbf{v}_{r+1}$ is the linear combination of the other vectors, a contradiction. Hence α_{r+1} must equals to zero. So the $r+1$ term in the equation (1.1) vanishes. And because the system $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ is linearly independent, all the scalars $\alpha_i = 0$ for all $i = 0, 1, \dots, r$. Thus, the system

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}$$

is also *linearly independent*. $\square$

Exercise 2.6 Is it possible that vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are linearly dependent, but the vectors $\mathbf{w_1} = \mathbf{v_1} + \mathbf{v_2}$, $\mathbf{w_2} = \mathbf{v_2} + \mathbf{v_3}$ and $\mathbf{w_3} = \mathbf{v_3} + \mathbf{v_1}$ are linearly *independent*.

Proof. It's not possible, and we're going to prove this assertion via contradiction. Assume that there are such vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ satisfying the above conditions. Then there are numbers $x, y, z \in \mathbb{R}$ such that

$$|x| + |y| + |z| > 0 \quad \text{and} \quad x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3 = \mathbf{0}.$$

By letting

$$a = x + y - z, \quad b = y + z - x, \quad c = z + x - y$$

we obtain that

$$\begin{aligned} a\mathbf{w}_1 + b\mathbf{w}_2 + c\mathbf{w}_3 &= (x\mathbf{w}_1 + y\mathbf{w}_1 - z\mathbf{w}_1) + (y\mathbf{w}_2 + z\mathbf{w}_2 - x\mathbf{w}_2) \\ &\quad + (x\mathbf{w}_3 + z\mathbf{w}_3 - y\mathbf{w}_3) \\ &= 2x\mathbf{v}_1 + 2y\mathbf{v}_2 + 2z\mathbf{v}_3 \\ &= \mathbf{0}. \end{aligned}$$

Since $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ are linearly independent, we must have $a = b = c = 0$. Hence

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

adding all the 3 eqations, $x + y + z = 0$. Substituting back to the system of eqations above we get

$$x = y = z = 0$$

which contradicts to the fact that $|x| + |y| + |z| > 0$. $\square$

Exercise 2.7 Any finite independent system is a subset of some basis.

Proof. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent. If this system is generating, then it's a base and we're done. If not, from exercise 2.5, there exists $\mathbf{v}_{n+1}$ such that

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}\}$$

is still linearly independent. Now if this new system is generating, then we're done. If not, we keep continue this process a finite steps, adding vectors $\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \ldots, \mathbf{v}_{n+r}$, and eventually the new system

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+r}\}$$

is now a basis. $\square$

3. Linear Transformation

Homework 1. Prove that the transformation $T : \mathbb{F}^n \to \mathbb{F}^m$ if and only if $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$ for any scalars α, β and vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}$.

Proof. We need to prove this in two directions.

($\Rightarrow$) Suppose T is a linear transformation, then

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) = T(\alpha\mathbf{x}) + T(\beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$$

as needed.

($\Leftarrow$) For this direction, we first assume that T has the property that $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$ for all $\alpha, \beta, \mathbf{x}, \mathbf{y}$. We need to show that T has the property listed in the definition of the linear transformation. Observe that

- take $\alpha = \beta = 1$ then, $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$
- take $\beta = 0$ then, $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$

Hence T is a linear transformation, and the proof is completed.

$\square$

Homework 2. Let $T : V \to W$ be a linear transformation. Prove that $T(\mathbf{0}) = \mathbf{0}$ and

$$TV = \{T\mathbf{v} \; : \; \mathbf{v} \in V\}$$

is a vector space.

Proof. Since T is linear, and as proved before $0 \cdot \mathbf{0} = \mathbf{0}$, it's easy to see that

$$T(\mathbf{0}) = T(0 \cdot \mathbf{0}) = 0 \cdot T(\mathbf{0}) = \mathbf{0}.$$

To prove that TV is a vector space, we need to check that TV satisfies all the eight conditions listed in the definition of vector space.

We first need to prove that TV is closed. Because $TV \subset W$, hence TV is closed under scalar multiplication and vector addition. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$. Observe that

- $T\mathbf{x} + T\mathbf{y} = T\mathbf{y} + T\mathbf{x} \quad$ (commutitativity of W)
- $(T\mathbf{x} + T\mathbf{y}) + T\mathbf{z} = T\mathbf{x} + (T\mathbf{y} + T\mathbf{z}) \quad$ (associativity of W)
- The vector $\mathbf{0} \in W$ is the indentity of TV because

$$T\mathbf{x} + \mathbf{0} = T\mathbf{x} + T\mathbf{0} = T(\mathbf{x} + \mathbf{0}) = T(\mathbf{x}), \quad \forall \mathbf{x} \in V$$

- The vector $T(-\mathbf{x})$ is the additive inverse of $T\mathbf{x}$ because

$$T\mathbf{x} + T(-\mathbf{x}) = T(\mathbf{x} - \mathbf{x}) = \mathbf{0}$$

- $1 \cdot T\mathbf{v} = T\mathbf{v} \quad$ (multiplicative iden. in W)

Let α, β be scalars.

- multiplicative associativity

$$\begin{aligned}
(\alpha\beta)T\mathbf{x} &= T((\alpha\beta)\mathbf{x}) && \text{(linearity of } T) \\
&= T(\alpha(\beta\mathbf{x})) && \text{(mult. asso. of } V) \\
&= \alpha T(\beta\mathbf{x}) && \text{(linearity of } T) \\
&= \alpha \cdot \beta T\mathbf{x}
\end{aligned}$$

- scalar multiplication

$$\begin{aligned}
\alpha(T\mathbf{x} + T\mathbf{y}) &= \alpha T(\mathbf{x} + \mathbf{y}) && \text{(linearity of } T) \\
&= T(\alpha(\mathbf{x} + \mathbf{y})) \\
&= T(\alpha\mathbf{x} + \alpha\mathbf{y}) && \text{(scalar mult. in } V) \\
&= T(\alpha\mathbf{x}) + T(\alpha\mathbf{x}) && \text{(linearity of } T) \\
&= \alpha T\mathbf{x} + \alpha T\mathbf{y}
\end{aligned}$$

- scalar multiplication

$$\begin{aligned}
(\alpha + \beta)T\mathbf{x} &= T((\alpha + \beta)\mathbf{x}) && \text{(linearity of } T) \\
&= T(\alpha\mathbf{x} + \beta\mathbf{x}) && \text{(scalar mult. of } V) \\
&= T(\alpha\mathbf{x}) + T(\beta\mathbf{x}) \\
&= \alpha T\mathbf{x} + \beta T\mathbf{x}
\end{aligned}$$

We see that TV has all eight properties to be a vector space, and the proof is completed.

$\square$

Homework 3. Let V, W be vector spaces. Prove that $\mathcal{L}(V, W)$, the set of all linear transformations $T : V \to W$, is also a vector space.

Proof. We first need to show that $\mathcal{L}(V, W)$ is closed. Let $T_1, T_2 \in \mathcal{L}(V, W)$ and a be a scalar. So we need to show the transformation $T_1 + T_2$ and aT_1 are both linear.

- Let $\mathbf{x}, \mathbf{y}$ be arbitrary vectors in V and α, β be scalar. Denote $T := T_1 + T_2$. Observe that

$$\begin{aligned} T(\alpha\mathbf{x} + \beta\mathbf{y}) =& (T_1 + T_2)(\alpha\mathbf{x} + \beta\mathbf{y}) \\ &= T_1(\alpha\mathbf{x} + \beta\mathbf{y}) + T_2(\alpha\mathbf{x} + \beta\mathbf{y}) && \text{(by def. of } T_1 + T_2) \\ &= \alpha T_1\mathbf{x} + \beta T_1\mathbf{y} + \alpha T_2\mathbf{x} + \beta T_2\mathbf{y} && \text{(by lin. of } T_1 \text{ and } T_2) \\ &= (\alpha T_1\mathbf{x} + \alpha T_2\mathbf{x}) + (\beta T_1\mathbf{y} + \beta T_2\mathbf{y}) \\ &= \alpha(T_1\mathbf{x} + T_2\mathbf{x}) + \beta(T_1\mathbf{y} + T_2\mathbf{y}) && \text{(by scalar mult. in } W) \\ &= \alpha(T_1 + T_2)\mathbf{x} + \beta(T_1 + T_2)\mathbf{y} \\ &= \alpha T\mathbf{x} + \beta T\mathbf{y} \end{aligned}$$

  This shows that $T_1 + T_2$ is also a linear transformation, hence $\mathcal{L}(V, W)$ is closed under addition.

- Similarly, we let $\mathbf{x}, \mathbf{y} \in V$. For simplicity, we again denote $T := aT_1$. Hence for any scalars α, β

$$\begin{aligned} T(\alpha\mathbf{x} + \beta\mathbf{y}) &= (aT_1)(\alpha\mathbf{x} + \beta\mathbf{y}) \\ &= a \cdot T_1(\alpha\mathbf{x} + \beta\mathbf{y}) && \text{(by def. of } aT_1) \\ &= a \cdot (\alpha T_1\mathbf{x} + \beta T_1\mathbf{y}) && \text{(by lin. of } T_1) \\ &= \alpha a T_1\mathbf{x} + \beta a T_1\mathbf{y} \\ &= \alpha(aT_1)\mathbf{x} + \beta(aT_1)\mathbf{y} \\ &= \alpha T\mathbf{x} + \beta T\mathbf{y} \end{aligned}$$

  This suggests that aT_1 is also linear, hence $\mathcal{L}(V, W)$ is closed under scalar multiplication. Ultimately, we've proved that $\mathcal{L}(V, W)$ is closed as needed.

We are now ready to prove that $\mathcal{L}(V, W)$ is a vector space. Let $T_1, T_2, T_3 \in \mathcal{L}(V, W)$ we have

- $T_1 + T_2 = T_2 + T_1$, because for any $\mathbf{x} \in V$

$$(T_1 + T_2)\mathbf{x} = T_1\mathbf{x} + T_2\mathbf{x} = T_2\mathbf{x} + T_1\mathbf{x} = (T_2 + T_1)\mathbf{x}.$$

- $T_1 + (T_2 + T_3) = (T_1 + T_2) + T_3$, because for any $\mathbf{x} \in V$

$$\begin{aligned} (T_1 + (T_2 + T_3))\mathbf{x} &= T_1\mathbf{x} + (T_2 + T_3)\mathbf{x} \\ &= T_1\mathbf{x} + (T_2\mathbf{x} + T_3\mathbf{x}) \\ &= (T_1\mathbf{x} + T_2\mathbf{x}) + T_3\mathbf{x} && \text{(by asso. of } W) \\ &= (T_1 + T_2)\mathbf{x} + T_3\mathbf{x} \\ &= ((T_1 + T_2) + T_3)\mathbf{x} \end{aligned}$$

- Consider the transformation $0 : V \to W$ such that $0(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v} \in V$. We're going to prove that this 0 is the indentity of $\mathcal{L}(V, W)$. But first, we need to know if 0 is linear or not. For any $\mathbf{v_1}, \mathbf{v_2} \in V$, we have

$$0(\alpha\mathbf{v_1} + \beta\mathbf{v_2}) = \mathbf{0} \quad \text{and} \quad \alpha 0\mathbf{v_1} + \beta 0\mathbf{v_2} = \alpha\mathbf{0} + \beta\mathbf{0} = \mathbf{0}.$$

  Hence $0(\alpha\mathbf{v_1} + \beta\mathbf{v_2}) = \alpha 0\mathbf{v_1} + \beta 0\mathbf{v_2}$, thus the transformation 0 is linear, i.e. $0 \in \mathcal{L}(V, W)$.

  Observe that for any $\mathbf{x} \in V$

$$(T_1 + 0)\mathbf{x} = T_1\mathbf{x} + 0\mathbf{x} = T_1\mathbf{x}$$

  This implies that $T_1 + 0 = T_1$ for any $T_1 \in \mathcal{L}(V, W)$. We conclude that 0 is the indentity of $\mathcal{L}(V, W)$.

- The transformation $-T_1 := (-1)T_1$ is the additive inverse of T_1 because for any $\mathbf{x} \in V$

$$T_1\mathbf{x} + (-T_1\mathbf{x}) = T_1\mathbf{x} + T_1(-\mathbf{x}) = T_1(\mathbf{x} - \mathbf{x}) = \mathbf{0} = 0(\mathbf{x}).$$

- $1 \cdot T_1 = T_1$ because $(1 \cdot T_1)\mathbf{x} = 1 \cdot T_1\mathbf{x} = T_1\mathbf{x}$ for any $\mathbf{x} \in V$.
- $(\alpha\beta)T_1 = \alpha(\beta T_1)$, because

$$[(\alpha\beta)T_1]\mathbf{x} = (\alpha\beta)T_1\mathbf{x} = T_1(\alpha\beta\mathbf{x}) = \alpha T_1(\beta\mathbf{x}) = \alpha(\beta T_1)\mathbf{x}$$

- $\alpha(T_1 + T_2) = \alpha T_1 + \alpha T_2$ because

$$[\alpha(T_1 + T_2)](\mathbf{x}) = \alpha T_1\mathbf{x} + \alpha T_2\mathbf{x} = (\alpha T_1 + \alpha T_2)(\mathbf{x})$$

$\square$

Exercise 3.1 Multiply

a). $$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 54 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1+6+6 \\ 4+15+12 \end{pmatrix} = \begin{pmatrix} 13 \\ 31 \end{pmatrix}$$

b). $$\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1+6 \\ 0+3 \\ 2+0 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix}$$

c). $$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1+4+0+0 \\ 0+2+6+0 \\ 0+0+3+8 \\ 0+0+0+4 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \\ 11 \\ 4 \end{pmatrix}$$

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

can't be multiplied because the number of columns of the first matrix doesn't equal to the number of rows of the second matrix.

Exercise 3.2 Let a linear transformation in $\mathbb{R}^2$ be the reflection in the line $x_1 = x_2$. Find its matrix.

Solution. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be this transformation. The basis of the domain is $\{\mathbf{e_1}, \mathbf{e_2}\}$ where $\mathbf{e_1} = (1,0)^T$ and $\mathbf{e_2} = (0,1)^T$. Because T reflect the line $x_1 = x_2$ then

$$T\mathbf{e_1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad T\mathbf{e_2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Therefore, the matrix of this transformation is $[T] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

□

Exercise 3.3 For each linear transformation below, find its matrix

a). $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x,y)^T = (x + 2y, 2x - 5y, 7y)^T$

b). $T : \mathbb{R}^4 \to \mathbb{R}^3$ defined by $T(x_1, x_2, x_3, x_4)^T = (x_1 + x_2 + x_3 + x_4, x_2 - x_4, x_1 + 3x_2 + 6x_4)^T$

c). $T : \mathbb{P}_n \to \mathbb{P}_n$ st $Tf(t) = f'(t)$ (find the matrix with respect to the standard basis $1, t, t^2, \ldots, t^n$)

d). $T : \mathbb{P}_n \to \mathbb{P}_n$ st $Tf(t) = 2f(t) + 3f'(t) - 4f''(t)$.

Proof. Find the matrix.

a). The standard basis in $\mathbb{R}^2$ is $\{\mathbf{e_1}, \mathbf{e_2}\}$ where $\mathbf{e_1} = (1,0)^T$ and $\mathbf{e_2} = (0,1)^T$. We have

$$T\mathbf{e_1} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad \text{and} \quad T\mathbf{e_2} = \begin{pmatrix} 2 \\ -5 \\ 7 \end{pmatrix}$$

Hence $\begin{pmatrix} 1 & 2 \\ 2 & -5 \\ 0 & 7 \end{pmatrix}$ is its matrix.

b). Let $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, \mathbf{e_4}\}$ be the standard basis in $\mathbb{R}^4$. Hence

$$T\mathbf{e_1} = T(1,0,0,0)^T = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad T\mathbf{e_2} = T(0,1,0,0)^T = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

$$T\mathbf{e_3} = T(0,0,1,0)^T = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad T\mathbf{e_4} = T(0,0,0,1)^T = \begin{pmatrix} 1 \\ -1 \\ 6 \end{pmatrix}$$

Therefore, $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 3 & 0 & 6 \end{pmatrix}$ is its matrix.

c). Let $E = \{t^n, t^{n-1}, \ldots, t, 1\}$ be the standard basis and $f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in \mathbb{P}_n$. We write

$$f(t) = (a_n, a_{n-1}, \ldots, a_1, a_0)^{\mathrm{T}}$$

is base E. Since

$$T(t^n) = nt^{n-1}, \quad T(t^{n-1}) = (n-1)t^{n-2} \quad , \ldots, \quad T(t) = 1, \quad T(1) = 0$$

Therefore its matrix is

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ n & 0 & 0 & \cdots & 0 & 0 \\ 0 & n-1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & n-2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

d). $Tf(t) = 2f(t) + 3f'(t) - 4f''(t)$
Again, the standard basis is $\{t^n, t^{n-1}, \ldots, t, 1\}$. For each $i \in [0, n]$ we have

$$T(t^i) = 2t^i + 3it^{i-1} - 4i(i-1)t^{i-2}$$

Hence the matrix is achieved by stacking $[T(t^n), \ldots, T(t^i), \ldots, T(t), T(1)]$, therefore the matrix is

$$[T] = \begin{bmatrix} 2 & 0 & \cdots & 0 & 0 \\ 3n & 2 & \cdots & 0 & 0 \\ -4n(n-1) & 3(n-1) & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 2 & 0 \\ 0 & 0 & \cdots & 3 & 2 \end{bmatrix}$$

□

Exercise 3.4 Find 3×3 matrices representing the transformations of $\mathbb{R}^3$ which

a). project every vector onto x-y plane;

b). reflect every vector through x-y plane;

c). rotate the x-y plane through $30°$, leaving the z-axis alone.

Proof. In space $\mathbb{R}^3$, we shall use its standard basis $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\}$ where $\mathbf{e_1} = (1,0,0)^{\mathrm{T}}$, $\mathbf{e_2} = (0,1,0)^{\mathrm{T}}$ and $\mathbf{e_3} = (0,0,1)^{\mathrm{T}}$.

a). Let T be this transformation. This means $T(x,y,z)^{\mathrm{T}} = (x,y,0)^{\mathrm{T}}$. We get

$$T\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad T\mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad T\mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

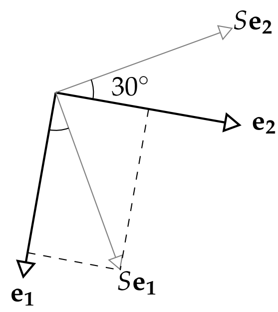
Therefore is matrix is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

b). Let R be this transformation. Since R project every vector through x-y plane, hence $R(x,y,z)^{\mathrm{T}} = (x,y,-z)^{\mathrm{T}}$. We get

$$R\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad R\mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad R\mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Thus the matrix of R is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

c). Let S be this transformation. S moves the vectors $\mathbf{e_1}, \mathbf{e_2}$ to the point x', y' respectively.



Since $\cos 30° = \frac{\sqrt{3}}{2}$ and $\sin 30° = \frac{1}{2}$, we conclude that

$$S\mathbf{e_1} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}, \quad S\mathbf{e_2} = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{pmatrix}, \quad S\mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Therefore the matrix is $\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

□

Exercise 3.5 Let A be a linear transformation. If $\mathbf{z}$ is the center of the staight interval $[\mathbf{x}, \mathbf{y}]$, show that $A\mathbf{z}$ is the center of the interval $[A\mathbf{x}, A\mathbf{y}]$.

Proof. $\mathbf{z}$ is the center of $[\mathbf{x}, \mathbf{y}]$ iff $\mathbf{z} = \frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}$. Therefore,

$$A\mathbf{z} = A\left(\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}\right) = \frac{1}{2}A\mathbf{x} + \frac{1}{2}A\mathbf{y}$$

Thus, $A\mathbf{z}$ is the center of the interval $[A\mathbf{x}, A\mathbf{y}]$. □

Exercise 3.6 The set $\mathbb{C}$ of complex numbers can be canonically identified with the space $\mathbb{R}^2$ by treating each $z = x + iy \in \mathbb{C}$ as a column $(x, y)^T \in \mathbb{R}^2$.

a). Treating $\mathbb{C}$ as a complex vector space, show that the multiplication by $\alpha = a + ib \in \mathbb{C}$ is a linear transformation in $\mathbb{C}$. What is its matrix.

b). Treating $\mathbb{C}$ as the real vector space $\mathbb{R}^2$ show that the multiplication by $\alpha = a + ib \in \mathbb{C}$ is a linear transformation there.

c). Define $T(x + iy) = 2x - y + i(x - 3y)$. Show that this tran is not a linear transformation in the complex vector space $\mathbb{C}$, but if we treat $\mathbb{C}$ as the real vector space $\mathbb{R}^2$ then it is a linear transformation there, then find its matrix.

Proof.

a). Let T be this transformation. For any $\mathbf{x} \in \mathbb{C}$, we have $T\mathbf{x} = \alpha\mathbf{x} \in \mathbb{C}$. Thus $T : \mathbb{C} \to \mathbb{C}$, and we'll prove that T is a linear transformation. Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}$ be two vectors, and $z \in \mathbb{C}$ be a scalar (complex). Observe that

- $T(\mathbf{x} + \mathbf{y}) = \alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y} = T\mathbf{x} + T\mathbf{y}$ (distributivity of complex numbers)
- $T(z\mathbf{x}) = \alpha(z\mathbf{x}) = z(\alpha\mathbf{x}) = zT\mathbf{x}$

This shows that this transformation T is a linear one. To find its matrix, we only need to know the basis of $\mathbb{C}$. Since any vector $\mathbf{x} \in \mathbb{C}$ we be written as

$$\mathbf{x} = 1 \cdot \underbrace{\mathbf{x}}_{\text{scalar}}$$

and because this representation is unique, we obtain that $\{1\} \subset \mathbb{C}$ is a basis of $\mathbb{C}$. Thus the matrix is

$$[T] = [T(1)] = [\alpha \cdot 1] = [\alpha].$$

b). Because we treat $\mathbb{C}$ as $\mathbb{R}^2$, then any complex number $\mathbf{x} = x + iy$ can be represented as $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$. Let T be this transformation. Thus T would look like

$$\begin{aligned} T\begin{pmatrix} x \\ y \end{pmatrix} &= T(\mathbf{x}) = \alpha\mathbf{x} \\ &= (a + ib)(x + iy) \\ &= (ax - by) + i(ay + bx) \\ &= \begin{pmatrix} ax - by \\ ay + bx \end{pmatrix} \in \mathbb{R}^2 \end{aligned}$$

Thus $T : \mathbb{R}^2 \to \mathbb{R}^2$. We need to show that T is in fact linear. Let $\mathbf{x_1} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\mathbf{x_2} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ and be two arbitrary vectors. We have

$$T\mathbf{x_1} + T\mathbf{x_2} = \begin{pmatrix} ax_1 - by_1 \\ ay_1 + bx_1 \end{pmatrix} + \begin{pmatrix} ax_2 - by_2 \\ ay_2 + bx_2 \end{pmatrix} = \begin{pmatrix} a(x_1 + x_2) - b(y_1 + y_2) \\ a(y_1 + y_2) + b(x_1 + x_2) \end{pmatrix} = T(\mathbf{x_1} + \mathbf{x_2}),$$

and for any scalar $r \in \mathbb{R}$,

$$rT\mathbf{x} = r\begin{pmatrix} ax - by \\ ay + bx \end{pmatrix} = \begin{pmatrix} rax - rby \\ ray + rbx \end{pmatrix} = T(r\mathbf{x}).$$

This shows that T is a linear transformation. To find the matrix, we first need to find a bisis in $\mathbb{R}^2$. Luckily, as we've proved earlier we could choose $\{\mathbf{e_1}, \mathbf{e_2}\}$ to be a basis where

$$\mathbf{e_1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{e_2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Therefore

$$T\mathbf{e_1} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{and} \quad T\mathbf{e_2} = \begin{pmatrix} -b \\ a \end{pmatrix}$$

Thus the matrix of this transformation is $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

c). Define $T(x+iy) = 2x - y + i(x - 3y)$

- We'll prove that T is not linear in complex vector space. Observe that

$$T(i) = T(0+i) = -1-3i \quad \text{and} \quad T(1) = T(1+0i) = 2+i$$

  cleary $T(i) \neq iT(1)$, this implies that T is not a linear transformation in $\mathbb{C}$.

- In $\mathbb{R}^2$ the transformation would look like

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ x - 3y \end{pmatrix}.$$

  For any vectors $\mathbf{x_1} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\mathbf{x_2} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, we have

$$T\mathbf{x_1} + T\mathbf{x_2} = \begin{pmatrix} 2x_1 - y_1 \\ x_1 - 3y_1 \end{pmatrix} + \begin{pmatrix} 2x_2 - y_2 \\ x_2 - 3y_2 \end{pmatrix} = \begin{pmatrix} 2(x_1+x_2) - (y_1+y_2) \\ (x_1+x_2) - 3(y_1+y_2) \end{pmatrix} = T(\mathbf{x_1} + \mathbf{x_2})$$

  and for any scalar $r \in \mathbb{R}$,

$$rT\mathbf{x} = r\begin{pmatrix} 2x - y \\ x - 3y \end{pmatrix} = \begin{pmatrix} 2rx - ry \\ rx - 3ry \end{pmatrix} = T(r\mathbf{x})$$

  this shows that $T : \mathbb{R}^2 \to \mathbb{R}^2$ is linear. Because $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is a basis in $\mathbb{R}^2$ and

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{and} \quad T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -3 \end{pmatrix},$$

  thus the matrix of this transformation is $\begin{pmatrix} 2 & -1 \\ 1 & -3 \end{pmatrix}$.

$\square$

Exercise 3.7 Show that any linear transformation in $\mathbb{C}$ (treated as a complex vector space) is a multiplication by $\alpha \in \mathbb{C}$.

Proof. Let $T : \mathbb{C} \to \mathbb{C}$ be this transformation. For any $\mathbf{x} \in \mathbb{C}$

$$T\mathbf{x} = T(\mathbf{x} \cdot \mathbf{1}) = \mathbf{x} \cdot \underbrace{T(\mathbf{1})}_{\text{scalar}}$$

and the proof is completed. $\square$

4. Linear transformation as a Vector

Let set $\mathcal{L}(V, W)$ is a vector space with addition and scalar multiplication (as proved above).

5. Composition

Homework 4. Let A and B be matrices of size $m \times n$ and $n \times m$ respectively. Then

$$\text{trace}(AB) = \text{trace}(BA).$$

Proof. We'd like to prove this theorem *less* computationally. Let $X \in M_{n\times m}$. Consider the mapping $T,\ T_1 : M_{n\times m} \to \mathbb{F}$ defined by

$$T(X) = \text{trace}(AX) \quad \text{and} \quad T_1(X) = \text{trace}(XA).$$

To prove the theorem it is sufficient to show that T, T_1 are linear and they are the same. so by substituting $X = B$ gives the theorem.

Claim 1. The transformations T, T_1 defined above are linear.

Proof. For $X, Y \in M_{n\times m}$,

- From the properties of matrix, $A(X+Y) = AX + AY$. Because AX and BX are both square matrices with size $m \times m$, and since we add the matrices $AX + AY$ entrywise, it follows that

$$\begin{aligned} T(X+Y) &= \text{trace}(A(X+Y)) = \text{trace}(AX+AY) \\ &= \text{trace}(AX) + \text{trace}(AY) \\ &= T(X) + T(Y) \end{aligned}$$

- Similarly for any scalar $\alpha \in \mathbb{F}$,

$$T(\alpha X) = \text{trace}(A \cdot \alpha X) = \text{trace}(\alpha AX) = \alpha\, \text{trace}(AX) = \alpha T(X)$$

This implies that T is a linear transformation. With simply proof, we conclude that T_1 is also a linear transformation. □

We choose $\mathbf{e_{11}}, \mathbf{e_{21}}, \dots, \mathbf{e_{nm}}$ to be the standard basis of $M_{n\times m}$, meaning the vector

$$\mathbf{e_{ij}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \cdots & & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

is a matrix whose entries are zero, except at the entry at row i and column j, which is 1. Then we only need to show that $T\mathbf{e_{ij}} = T_1\mathbf{e_{ij}}$ for all i, j. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} \\ \vdots & & \vdots & & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{im} \\ a_{nn} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nm} \end{pmatrix}$$

Hence

$$A\mathbf{e_{ij}} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & a_{ij} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \cdot & \cdot \\ \cdot & a_{ij} & \cdot \\ \cdot & \cdot & 0 \end{pmatrix}$$

and

$$\mathbf{e_{ij}}A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & a_{ij} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \begin{pmatrix} 0 & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot \\ \cdot & \cdot & a_{ij} & \cdot \\ \cdot & \cdot & \cdot & 0 \end{pmatrix}$$

This implies that $T\mathbf{e_{ij}} = T_1\mathbf{e_{ij}}$ for all i, j, and hence $T = T_1$. □

Exercise 5.1 Working on it.

Exercise 5.2 Let T_γ be the rotation matrix by γ in $\mathbb{R}^2$. Check by matrix multiplication that $T_\gamma T_{-\gamma} = T_{-\gamma}T_\gamma = I$.

Proof. Working on it. □

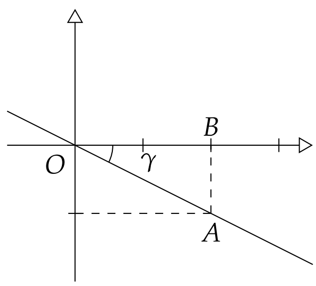
Exercise 5.3 Multiply two rotation matrices T_α and T_β. Deduce formulas for $\sin(\alpha+\beta)$ and $\cos(\alpha+\beta)$ from here.

Proof. Working on it. □

Exercise 5.4 Find the matrix of the orthogonal projection in $\mathbb{R}^2$ on to the line $x_1 = -2x_2$.

Proof. Let T be this transformation. Let R_γ and P_x be the transformations of rotation by γ and projection to x-axis, respectively. Therefore $T = R_\gamma P_x R_{-\gamma}$. Note that

$$R_\gamma = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}, \quad R_{-\gamma} = \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix}, \quad \text{and} \quad P_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$



Thus,

$$\cos\gamma = \frac{\overline{OB}}{\overline{OA}} = \frac{2}{\sqrt{5}} \quad \text{and} \quad \sin\gamma = \frac{\overline{AB}}{\overline{OA}} = \frac{-1}{\sqrt{5}}.$$

We get

$$
\begin{aligned}
T &= \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} 4 & -2 \\ -2 & -2 \end{pmatrix}
\end{aligned}
$$

□

Exercise 5.5 Find linear transformations $A, B : \mathbb{R}^2 \to \mathbb{R}^2$ such that $AB = \mathbf{0}$ but $BA \neq \mathbf{0}$.

Solution. Consider the following

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

however

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \neq \mathbf{0}$$

Therefore, these two matrices are the ones we wish to find. □

Exercise 5.6 Prove that $\operatorname{trace}(AB) = \operatorname{trace}(BA)$.

Proof. See on page 18. □

Exercise 5.7 Construct a non-zero matrix A such that $A^2 = \mathbf{0}$

Proof. Let $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Thus perform the multiplication, we get

$$\begin{cases} a^2 + bc = 0 \\ ac + cd = 0 \\ ab + bd = 0 \\ bc + d^2 = 0 \end{cases}$$

for simplicity, we'll choose $a = 1$. Hence $bc = -1$ and

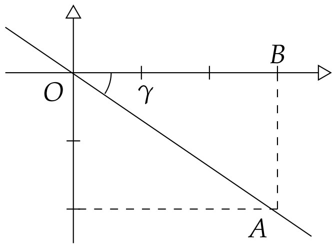
$$\begin{cases} c(d+1) = 0 \\ b(d+1) = 1 \\ d^2 = 1 \end{cases}$$

this suggests that $d = -1$, and $bc = -1$. Here, we'll choose $b = 1$ and $c = -1$. Therefore, the matrix

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.$$

□

Exercise 5.8 Find the matrix of the reflection through the line $y = -2x/3$.



Proof. Let T be this transformation and γ be the angle between the x-axis and the line $y = -2x/3$. Hence $T = R_\gamma T_0 R_\gamma$. We then have $\cos\gamma = OB/OA = 3/\sqrt{13}$ and $\sin\gamma = -AB/OA = -2/\sqrt{13}$. Thus

$$\begin{aligned} T &= \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix} \\ &= \frac{1}{13} \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \\ &= \frac{1}{13} \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ -2 & -3 \end{pmatrix} \\ &= \frac{1}{15} \begin{pmatrix} 5 & -12 \\ -12 & -5 \end{pmatrix}. \end{aligned}$$

□

6. Isomophism

Exercise 6.1 Prove that if $A : V \to W$ is an isomorphism and $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ is a basis in V, then $A\mathbf{v_1}, A\mathbf{v_2}, \ldots, A\mathbf{v_n}$ is a basis in W.

Proof. Since $A : V \to W$ is an isomorphism, hence it's invertable i.e. there is a linear transformation $A^{-1} : W \to V$ such that $AA^{-1} = A^{-1}A = I$. Thus for any $\mathbf{w} \in W$, there is a $\mathbf{v} \in V$ such that $A^{-1}\mathbf{w} = \mathbf{v}$. Recall that $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis in V, then there are unique scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\mathbf{v} = \alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2} + \cdots + \alpha_n\mathbf{v_n}.$$

This implies that

$$\begin{aligned} A^{-1}\mathbf{w} &= \alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2} + \cdots + \alpha_n\mathbf{v_n} \\ AA^{-1}\mathbf{w} &= A(\alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2} + \cdots + \alpha_n\mathbf{v_n}) \\ \mathbf{w} &= \alpha_1 A\mathbf{v_1} + \alpha_2 A\mathbf{v_2} + \cdots + \alpha_n A\mathbf{v_n} \end{aligned}$$

Because $\alpha_1, \alpha_2, \cdots, \alpha_n$ are unique, we conclude that any $\mathbf{w} \in W$ can be represented as a unique linear combination of $A\mathbf{v_1}, A\mathbf{v_2}, \ldots, A\mathbf{v_2}$. Thus the proof is completed. □

Exercise 6.2 Find all right inverses of the 1×2 matrix (row) $A = (1, 1)$. Conclude from here thaat the row A is not left invertable.

Exercise 6.3 Find all the left inverses of the column $(1,2,3)^{\mathrm{T}}$.

Proof. Let $A = (1,2,3)^{\mathrm{T}}$. Because A is a 3×1 matrix, then its inverse, say B is a 1×3 matrix. Let $B = \begin{pmatrix} x & y & z \end{pmatrix}$. Hence

$$AB = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix}$$

This implies that $x + 2y + 3z = 1$ or $x = 1 - 2y - 3z$. Thus all the left inverses of A is in the form

$$B = \begin{pmatrix} 1 - 2y - 3z & y & z \end{pmatrix}$$

where y, z are arbitrary real numbers. □

Exercise 6.4 Is the column $(1,2,3)^{\mathrm{T}}$ right invertable?

Solution. The column $(1,2,3)^{\mathrm{T}}$ is not right invertable, because as proved in previous exercise the column $(1,2,3)^{\mathrm{T}}$ has more than one left inverses. □

Exercise 6.5 Find two matrices A and B that AB is invertable, but A and B are not.

Solution. Consider: $A = \begin{pmatrix} 2 & 1 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^{\mathrm{T}}$ Note that

$$AB = \begin{pmatrix} 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 + 2 - 3 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix}$$

However, as proved in previous exercise, we know that the matrix A is not invertable. And we wish to prove that B is not invertable either. To achieved this, we have to find two matrices that are right invertable to B. Observe that

$$\begin{pmatrix} 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix}$$

This suggests that B is not invertable. Therefore, we've fond matrices A and B such that AB is invertable, yet A and B are not. □

Exercise 6.6 Suppose the product AB is invertable. Show that A is right invertable, and B is left invertable.

Proof. Because AB is invertable, then matrix $(AB)^{-1}$ is defined. Observe that

$$A \cdot B(AB)^{-1} = AB \cdot (AB)^{-1} = I$$

and

$$(AB)^{-1}A \cdot B = (AB)^{-1} \cdot AB = I$$

This shows that A is right invertable, and B is left invertable, as expected. □

Exercise 6.7 Let A and AB be invertable. Prove that B is also in invertable.

Proof. We claim that $(AB)^{-1}A$ is the inverse of B. To prove this, observe that

$$(AB)^{-1}A \cdot B = (AB)^{-1} \cdot AB = I$$

and

$$B \cdot (AB)^{-1}A = A^{-1}AB \cdot (AB)^{-1}A = A^{-1}IA = I$$

This shows that $(AB)^{-1}A$ is both the left and the right inverses of B. Thus, B is invertable. □

Exercise 6.8 Let A be $n \times n$ matrix. Prove that if $A^2 = \mathbf{0}$ then A is not invertable.

Proof. Assume by contradiction that A is invertable, meaning there's a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$. Thus

$$A = A^2A^{-1} = \mathbf{0}A^{-1} = \mathbf{0}$$

Then $I = AA^{-1} = \mathbf{0}A^{-1} = \mathbf{0}$, a contradiction. Therefore, A is not invertable. □

Exercise 6.9 Suppose $AB = \mathbf{0}$ for some non-zero matrix B. Can A invertable?

Proof. We claim that A is not invertable. To prove this, we assume by contradiction that A is invertable. Hence A^{-1} exists, and

$$B = A^{-1}A \cdot B = A^{-1} \cdot AB = A^{-1}\mathbf{0} = \mathbf{0}$$

which is a contradiction that B is non-zero. □

Exercise 6.10 Write matrices of the linear transformations T_1 and T_2 in $\mathbb{F}^5$, defined as follows: T_1 interchanges the coordinates X_2 and x_2 of the vector $\mathbf{x}$, and T_2 just adds to the coordinates $x_2 \to a$ times the coordinate x_4. and does not change other coordinates, i.e.

$$T_1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_4 \\ x_3 \\ x_2 \\ x_5 \end{pmatrix} \quad \text{and} \quad T_2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 + ax_4 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}.$$

where a is a fixed number. Show that T_1 and T_2 are invertable transformations, and write the matrices of the inverses.

Proof. The matrix of this □

7. Subspaces

Exercise 7.1 Let X and Y be subspaces of a vector space V. Prove that $X \cap Y$ is a subspace of V.

Proof. Let $\mathbf{a}$ and $\mathbf{b}$ be arbitrary vectors of $X \cap Y$. Because X and Y are themselves subspaces of V, hence

$$\begin{cases} \mathbf{a} \in X \\ \mathbf{b} \in X \end{cases} \implies \begin{cases} \alpha\mathbf{a} \in X \\ \beta\mathbf{b} \in X \end{cases}$$

this implies that $\alpha\mathbf{a} + \beta\mathbf{b} \in X$ for any scalars α, β. Similarly, $\alpha\mathbf{a} + \beta\mathbf{b} \in Y$. Thus $\alpha\mathbf{a} + \beta\mathbf{b} \in X \cap Y$. Therefore, $X \cap Y$ is also a subspace of V. $\square$

Exercise 7.2 Let V be a vector space. For $X, Y \subset V$ the sum $X + Y$ is the collection of all vectors $\mathbf{v}$ which can be represented as $\mathbf{v} = \mathbf{x} + \mathbf{y}$ where $\mathbf{x} \in X$ and $\mathbf{y} \in Y$. Show that if X and Y are subspaces of V, then $X + Y$ is also a subspace of V.

Proof. Let $\mathbf{v_1}, \mathbf{v_2} \in X + Y$ then there are $\mathbf{x_1}, \mathbf{x_2} \in X$ and $\mathbf{y_1}, \mathbf{y_2} \in Y$ such that

$$\mathbf{v_1} = \mathbf{x_1} + \mathbf{y_1} \quad \text{and} \quad \mathbf{v_2} = \mathbf{x_2} + \mathbf{y_2}$$

Because X, Y are subspaces of V, then

$$\alpha\mathbf{v_1} + \beta\mathbf{v_2} = \underbrace{(\alpha\mathbf{x_1} + \beta\mathbf{x_2})}_{\text{vector of } X} + \underbrace{(\alpha\mathbf{y_1} + \beta\mathbf{y_2})}_{\text{vector of } Y}$$

Hence $\alpha\mathbf{v_1} + \beta\mathbf{v_2}$ is also a vector of $X + Y$. Thus $X + Y$ is a subspace of V. $\square$

Exercise 7.3 Let X be a subspace of a vector space V, and let $\mathbf{v} \in V$ and $\mathbf{v} \notin X$. Prove that if $\mathbf{x} \in X$ then $\mathbf{x} + \mathbf{v} \notin X$.

Proof. We'll prove this by contradiction by assuming that $\mathbf{x} + \mathbf{v} \in X$. Because X is a subspace of V and $\mathbf{x} \in X$, hence $-\mathbf{x} \in X$. Thus

$$(\mathbf{x} + \mathbf{v}) + (-\mathbf{x}) \in X$$

and we conclude that $\mathbf{v} \in X$, which is a contradiction. $\square$

Exercise 7.4 Let X and Y be subspaces of a vector space V. Using the previous exercise, show that $X \cup Y$ is a subspace iff $X \subset Y$ or $Y \subset X$.

Proof. We need to prove this in two directions.

- If $X \subset Y$ or $Y \subset X$: Without loss of generality, we may assume that $Y \subset X$. Therefore $X \cup Y = X$ is a subspace of V.

- If $X \not\subset Y$ and $Y \not\subset Y$, we now wish to show that $X \cup Y$ is not a subspace. Observe that if $X \not\subset Y$ and $Y \not\subset X$ that means there are $\mathbf{x_0} \in X$ and $\mathbf{y_0} \in Y$ such that $\mathbf{x_0} \notin Y$ and $\mathbf{y_0} \notin X$. Hence $\mathbf{x_0}, \mathbf{y_0} \in X \cup Y$ and follow from the previous exercise, we conclude that

$$\mathbf{x_0} + \mathbf{y_0} \notin X \quad \text{and} \quad \mathbf{x_0} + \mathbf{y_0} \notin Y$$

  Therefore, $\mathbf{x_0} + \mathbf{y_0} \notin X \cup Y$. This suggests that $X \cup Y$ is not a vector space. Hence the proof is completed.

$\square$

<==

Chapter 2

Systems of linear equations

1. Different faces of linear transformation

2. Solution of a linear system. Echelon forms

Exercise 2.1 Write the systems of equations below in matrix form.

Exercise 2.2 Find all solutions of the vector equation

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + x_3\mathbf{v_3} = \mathbf{0}$$

where $\mathbf{v_1} = (1,1,0)^{\mathrm{T}}$, $\mathbf{x_2} = (0,1,1)^{\mathrm{T}}$ and $\mathbf{v_3} = (1,0,1)^{\mathrm{T}}$. What conclusion can you make about linear independence (dependence) of the system of vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$.

Proof. The echelon form of the system is

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

so, clearly the solution to this equation is $x_1 = x_2 = x_3 = 0$.

Conclusion. If the vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are linearly independence, then the above equation has unique solution, namely $x_1 = x_2 = x_3 = 0$.

If the vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are linearly dependence, then there exists α, β, γ (some of them are non-zero) such that

$$\alpha\mathbf{v_1} + \beta\mathbf{v_2} + \gamma\mathbf{v_3} = \mathbf{0}$$

Therefore, the solution to the above equation is

$$\begin{cases} x_1 = \alpha t \\ x_2 = \beta t \\ x_3 = \gamma t \end{cases}$$

for some $t \in \mathbb{R}$. $\square$

3. Analyzing pivots

Exercise 3.6 Prove or disprove. If the columns of a square $(n \times n)$ matrix are linearly independence, so are the columns of A^2.

Proof. Since A is a square matrix, and columns vectors are linearly independence, so

$$\begin{aligned} & \text{has pivot in every column} \\ \Longrightarrow\ & \text{has pivot in evry row} \\ \Longrightarrow\ & A \text{ is invertable} \\ \Longrightarrow\ & A^2 \text{ also invertable} \\ \Longrightarrow\ & A^2 \text{ has pivot every column and row} \end{aligned}$$

Therefore, the column vectors of A^2 also independence. □

Exercise 3.7 Prove or disprove. If the columns of a square $(n \times n)$ matrix are linearly independence, so are the columns of A^3.

Exercise 3.8 Show that if the equation $A\mathbf{x} = \mathbf{0}$ has unique solution, then A is left invertable.

Proof. Because A has unique solution, then it has pivot in every column. Therfore, #col $\leq$ #row, so we let $m \times n$ be the size of A where $n \leq m$ Let R be the reduced echelon form of A, hence there exists $E = E_k \cdots E_2 E_1$ such that $R = EA$. Observe that, R would look like

$$R = \begin{pmatrix} \boxed{1} & 0 & \cdots & 0 \\ 0 & \boxed{1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \boxed{1} \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

Using matrix muliplication gives us $R^{\mathrm{T}} R = I_n$. We obtain that

$$R^{\mathrm{T}} EA = R^{\mathrm{T}} T = I_n$$

Therfore, $\boxed{A \text{ is left invertable, and } R^{\mathrm{T}} E \text{ is its left inverse}}$.

□

4. Find A^{-1} by row reduction

5. Dimension

Exercise 5.1 True or false.

a). Every vector space that is generated by a finite set has a basis;

b). Every vector space has a (finite) basis;

c). A vector space cannot have more that one basis;

d). A vector space has a finite basis, then the number of vectors in every basis is the same;

e). The dimension of $\mathbb{P}_n$ is n;

f). The dimension of $M_{m\times n}$ is $m+n$;

g). If vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ generate (span) the vector space V, then every vector in V can be written as a linear combination of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ in only one way;

h). Every subspace of a finite-dimensional space is finite-dimensional.

i). If V is a vector space having dimension n, then V has exactly one subspace of dimension 0, and exactly one subspace of dimension n.

Proof. a). **True.** That finite set which generated a vector space is the spanning set itself. Since it's finite, it contains a basis.

b). **False.** Take $\mathbb{R}[x]$ for example.

c). **False.** In $\mathbb{R}^2$, one can choose

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} \quad \text{or} \quad \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$$

as a basis.

d). **True.** As proved in the above theorems,

$$\begin{aligned} &\#\,\text{independence vectors} \leq \dim V \\ &\#\,\text{generating vectors} \geq \dim V \end{aligned}$$

hence the number of any basis in V must be exactly $\dim V$ vectors.

e). **False.** In $\mathbb{P}_n$, the standard basis is

$$1,\ t,\ t^2, \ldots, t^n$$

which has $n+1$ vectors. Hence $\boxed{\dim \mathbb{P}_n = n+1}$.

f). **False.** The standard basis in $M_{m\times n}$ is

$$\{\mathbf{e}_{11}, \mathbf{e}_{12}, \ldots, \mathbf{e}_{mn}\}$$

has $m\times n$ vectors. Hence $\boxed{\dim M_{m\times n} = mn}$.

g). **False.** span doesn't guarantee uniqueness.

h). **True.** Let W be a subspace of V. Because dim V finite, we can find

$$\mathcal{A} = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\} \subset V$$

that spans V. WLOG, we assume that none of vectors in $\mathcal{A}$ belongs to W (the unluckiest case.) We can choose $\mathbf{w_1} \in W$ such that $\mathbf{w_1} \neq \mathbf{0}$. Then

$$\mathbf{w_1} = \alpha_1 \mathbf{v_1} + \cdots + \alpha_n \mathbf{v_n}$$

Because $\mathbf{w_1} \neq 0$, we're sure some of the α_i's are non-zero, say α_1. Then the new system

$$\mathcal{A}_1 = \{\mathbf{w_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$$

still spans the space V. Now, if $\mathcal{B}_1 := \{\mathbf{w_1}\}$ doesn't span W, we can repeat the above procedure and find $\mathbf{w_2}$. We can do this at most n times, because once we reach the nth step, we have the new system $\mathcal{A}_n \subset W$ that spans the whole space V.

Therfore, after some finite $k \leq n$ step, we have

$$\mathcal{B}_k = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_k}\} \subset W$$

spans W. Hence, W is finite dimensional.

i). **Not sure.**

□

Exercise 5.2 Prove that if V is a vector space having dimension n, then a system of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ in V is linearly independent iff it spans V.

Proof.

($\Longrightarrow$) Suppose that $v_1, \ldots, v_n$ linearly independent in V. If it's not span, then we can complete this system to a basis, but then this new system will have the number of vectors more than n, but this is a contradiction because a basis (linearly independent) cannot have vectors more than n.

($\Longleftarrow$) Now suppose that $v_1, \ldots, v_n$ spans V. If it's not linearly independence, we can throw away some vectors, and become the basis. However, this new basis system has less that n vectors in it, which is a contradiction because a basis (span) must have more than n vectors in it.

□

Exercise 5.3 Prove that a linearly independent system of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ in a vector space V is a basis iff $n = \dim V$.

Proof. We have $\mathbf{v_1}, \ldots, \mathbf{v_n}$ linearly independent. If it's a basis then $\dim V = n$. Now suppose that $\dim V = n$. We need to show that the system is a basis. From the above exercise, this system has to span V. Thus a basis in V. □

Exercise 5.4 (An old problem revisited.) Is it possible that vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent, but the vectors $\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{v}_2,\ \mathbf{w}_2 = \mathbf{v}_2 + \mathbf{v}_3$ and $\mathbf{w}_3 = \mathbf{v}_3 + \mathbf{v}_1$ are linearly independent.

Proof. Because $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ linearly independent, then

$$\dim \operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\} = 3.$$

Because $\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{v}_2,\ \mathbf{w}_2 = \mathbf{v}_2 + \mathbf{v}_3$, it's easy to see that

$$\operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\} \subseteq \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$$

But since the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ linearly dependent, then one vector, say $\mathbf{v}_3$, is the linear combination of $\mathbf{v}_1, \mathbf{v}_2$. This suggests that $\dim \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \dim \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\} \leq 2$, thus

$$\dim \operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\} \leq \dim \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \leq 2$$

But this contradicts to the result we found above. Hence it's impossible that such $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ exists. □

Exercise 5.5 Let vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be a basis in V. Show that $\mathbf{u} + \mathbf{v} + \mathbf{w},\ \mathbf{v} + \mathbf{w},\ \mathbf{w}$ is also a basis in V.

Proof. It can be shown by using the fact from the exercise above i.e. show that the new system is either spans or LI (because it already has 3 vectors in it). But here I'm gonna use matrix and pivot stuff. Since V has three vectors as a basis, then $\dim V = 3$. Choose an isomorphism $A : V \to \mathbb{R}^3$ such that $A\mathbf{v} = (1, 0, 0),\ A\mathbf{u} = (0, 1, 0)$ and $A\mathbf{w} = (0, 0, 1)$. Letting

$$\begin{aligned}
\mathbf{a}_1 &:= A(\mathbf{u} + \mathbf{v} + \mathbf{w}) = (1, 1, 1) \\
\mathbf{a}_2 &:= A(\mathbf{v} + \mathbf{w}) = (0, 1, 1) \\
\mathbf{a}_3 &:= A(\mathbf{w}) = (0, 0, 1)
\end{aligned}$$

Observe that the Echelon form of the matrix formed by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

has pivots in every row and every column, hence the system $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ is basis in $\mathbb{R}^3$. Using isomorphism theorem, we conclude that

$$\mathbf{u} + \mathbf{v} + \mathbf{w},\ \mathbf{v} + \mathbf{w},\ \mathbf{w}$$

also a basis in V. □

6. Rank

Exercise 6.4 Prove that if $A : X \to Y$ and V is a subspace of X, then $\dim AV \leq \operatorname{rank} A$. Then deduce that $\operatorname{rank} AB \leq \operatorname{rank} A$.

Proof. Observe that $AV \subseteq AX = \operatorname{col} A$, hence

$$\dim AV \leq \dim \operatorname{col} A = \operatorname{rank} A.$$

To prove the next statement, we may assume that $B : Z \to X$. This implies that $\operatorname{col} B = BZ \subseteq X$. Using the previous result we obtain that

$$\begin{aligned} & \dim A(\operatorname{col} B) \leq \operatorname{rank} A \\ \implies \quad & \dim \operatorname{col} AB \leq \operatorname{rank} A \\ \implies \quad & \operatorname{rank} AB \leq \operatorname{rank} A \end{aligned}$$

□

7. Change of basis

Exercise 7.3 Find the change of coordinates matrix that changes the coordinates in basis $\{1, 1+t\}$ in $\mathbb{P}_1$ to the coordinates in the basis $\{1-t, 2t\}$.

Proof. Let's denote $\mathcal{A} = \{1, 1+t\}$ and $\mathcal{B} = \{1-t, 2t\}$. Let $\mathcal{S}$ be the standard basis in $\mathbb{P}_1$. Therefore, the matrix that transforms from vector in basis $\mathcal{A}$ to basis $\mathcal{B}$ is $[\mathcal{B}\mathcal{A}] = [\mathcal{B}\mathcal{S}][\mathcal{S}\mathcal{A}]$. We have

$$[\mathcal{S}\mathcal{A}] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and

$$[\mathcal{B}\mathcal{S}] = [\mathcal{S}\mathcal{B}]^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

Therefore,

$$[\mathcal{B}\mathcal{A}] = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

□

Exercise 7.4 Let T be the ...

Proof. In standard basis, T looks like

$$[T] = \begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix}.$$

Let $\mathcal{B} = \{(1,1)^{\mathrm{T}}, (1,2)^{\mathrm{T}}\}$. In basis $\mathcal{B}$, the transformation would look like

$$[T]_{\mathcal{B}\mathcal{B}} = [\mathcal{B}\mathcal{S}][T][\mathcal{S}\mathcal{B}]$$

And we have

$$[\mathcal{S}\mathcal{B}] = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

so

$$[\mathcal{B}\mathcal{S}] = [\mathcal{S}\mathcal{B}]^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

Therefore, the transformation $[T]_{\mathcal{B}\mathcal{B}}$ in basis $\mathcal{B}$ is

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

□

Exercise 7.5 Prove that if A and B are similar matrices, then trace A = trace B.

Proof. Because A and B are similar, then there exists an invertable matrix Q such that $A = Q^{-1}BQ$. Observe that

$$A = Q^{-1} \cdot BQ \quad \text{and} \quad B = BQ \cdot Q^{-1}$$

This implies that trace A = trace B. □

Chapter 3

Determinant

3. Constructing the determinant

Exercise 3.4 A square $(n \times n)$ matrix is called skew-symmetric (or anti-symmetric) if $A^{\mathrm{T}} = -A$. Prove that if A is skew-symmetric and n is odd, then $\det A = 0$. Is it true for even n?

Proof. When n is odd, we have

$$\det A = \det A^{\mathrm{T}} = \det(-A) = (-1)^n \det A = -\det A$$

This implies that $\det A = 0$ whenever n is odd. For even n, the determinant $\det A$ doesn't neccessarily zero, for instance

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

is a skew-symmetric matrix, yet $\det B = 1$. □

Exercise 3.5 A square matrix is called *nilpoten* if $A^k = 0$ for some $k \in \mathbb{N}$. Show that if A is nilpoten, then $\det A = 0$.

Proof. We use the property of determinant,

$$\det 0 = \det A^k = (\det A)^k$$

This shows that $\det A = 0$. □

Exercise 3.6 Prove that if the matrices A and B are similar, then $\det A = \det B$.

Proof. Since $A \sim B$, hence there's invertable Q such that $A = QBQ^{-1}$. Therefore

$$\begin{aligned} \det A = \det QBQ^{-1} &= (\det Q)(\det B)(\det Q^{-1}) \\ &= \det(Q)(\det Q^{-1})(\det B) = \det(QQ^{-1})(\det B) \\ &= \det B. \end{aligned}$$

□

Exercise 3.7 A real square matrix Q is called *othogonal* if $Q^{\mathrm{T}}Q = I$. Prove that if Q is a n orthogonal matrix then $\det Q = \pm 1$.

Proof. We have
$$1 = \det Q^{\mathrm{T}}Q = (\det Q^{\mathrm{T}})(\det Q) = (\det Q)^2$$
this implies that $\det Q = \pm 1$. □

4. Formal definition

Exercise 4.2 Let P be a permutation matrix.

- Can you describe the corresponding linear transformation?
- Show that P is invertable. Can you describe P^{-1}.
- Show that for some $N > 0$, $P^N = I$.

Proof. Suppose that P is an $n \times n$ matrix.

- The linear transformation looks like it's swapping the axis in $\mathbb{R}^n$.
- If we interchange columns of P, we'll get the indentity matrix hence $\det P = \pm 1$. The direct computation shows that $P^{-1} = P^{\mathrm{T}}$.
- Because there are finitely many permutations, the sequence $\{P^n\}$ will eventually have repetitions. Hence we're sure there is i, j with $i > j$ such that $P^i = P^j$. Since P is invertable, we can multiply both sides by P^{-j}. Therefore
$$P^{i-j} = I,$$
now choose $N := i - j$ and this completes the proof.

□

Exercise 4.3 Why is there an even number of permutations of $(1, 2, \ldots, 9)$ and why are exactly half of them odd permutations?

Proof. There are 9! permutations of $(1, 2, \ldots, 9)$, which is an even number. To prove that there are exactly half of them odd (and other half even) we consider a 9×9 matrix A with all the entries are 1. The determinant is $\det A = 0$. However,
$$\det A = \sum_{\sigma \in \mathrm{perm}(9)} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \operatorname{sign}(\sigma)$$
Since all the $a_{jk} = 1$, we get
$$\sum_{\sigma} \operatorname{sign} \sigma = \det A = 0$$
Because there are even number of permutations, and $\operatorname{sign} \sigma = \pm 1$ we must have half of them has sign $= 1$ and the other half has sign $= -1$. □

5. Cofactor

6. Minor and rank

Exercise 6.2 Let A be an $n \times n$ matrix. How are $\det(3A)$, $\det(-A)$ and $\det(A^2)$ related to $\det A$?

Proof. It follows immediately that

$$\begin{aligned}\det(-A) &= (-1)^n \det A,\\ \det(3A) &= 3^n \det A,\\ \det(A^2) &= (\det A)^2.\end{aligned}$$

□

Exercise 6.3 If the entries of both A and A^{-1} are integers, is it possible that $\det A = 3$?

Proof. It's impossible. If it was such a case, we'll get $\det A^{-1} = 1/\det A = 1/3$. Since all the entries of A^{-1} are integers, so

$$\det A^{-1} = \sum_{\sigma} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \operatorname{sign} \sigma \in \mathbb{Z}$$

which is a contradiction that $\det A^{-1} = 1/3$. □

Exercise 6.4 Let $\mathbf{v_1}, \mathbf{v_2}$ be vectors in $\mathbb{R}^2$ and let A be the 2×2 matrix with columns $\mathbf{v_1}, \mathbf{v_2}$. Prove that $|\det A|$ is the area of parallelogram with two sides given by vectors $\mathbf{v_1}$ and $\mathbf{v_2}$.

Proof. Let α be angle between $\mathbf{v_1}$ and the x-axis, and let R be the matrix of rotation by $-\alpha$ angle. Denote

$$\tilde{\mathbf{v}}_\mathbf{1} := R\mathbf{v_1} = (a, 0) \quad \text{and} \quad \tilde{\mathbf{v}}_\mathbf{2} := R\mathbf{v_2} = (b, c).$$

for some reals a, b, c. Note that after the transformation, the area stays the same, i.e. $\text{area}(\mathbf{v_1}, \mathbf{v_2}) = \text{area}(\tilde{\mathbf{v}}_\mathbf{1}, \tilde{\mathbf{v}}_\mathbf{2})$. It's easy to see that the area of the new parallelogram is

$$|ac| = |\det(\tilde{\mathbf{v}}_\mathbf{1}, \tilde{\mathbf{v}}_\mathbf{2})|$$

(it has base $|a|$, and height $|c|$). But

$$\det(\tilde{\mathbf{v}}_\mathbf{1}, \tilde{\mathbf{v}}_\mathbf{2}) = \det(R\mathbf{v_1}, R\mathbf{v_2}) = \det(RA) = \det A$$

because $\det R = 1$ (rotation matrix). This implies that the area of the parallelogram is $|\det A|$. □

Exercise 6.5 Let $\mathbf{v_1}, \mathbf{v_2}$ be vectors in $\mathbb{R}^2$. Show that $\det(\mathbf{v_1}, \mathbf{v_2}) > 0$ if and only if there's a rotation T_α such that $T_\alpha \mathbf{v_1}$ parallel to $\mathbf{e_1}$ and $T_\alpha \mathbf{v_2}$ is in the upper half-plane.

Proof.

($\Rightarrow$) For this direction, the proof is almost identical to the previous exercise. We claim that $T_\alpha = R$ (R defined in the previous exercise). We now wanna show that $\tilde{\mathbf{v}}_2 = (b, c)$ is in the upper-half plane, meaning $c > 0$. But this immediately true from the fact that $a > 0$ and

$$0 < \det(\mathbf{v_1}, \mathbf{v_2}) = \det(\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2) = ac$$

($\Leftarrow$)

□

<==

Chapter 4

Intro to Spectral Theory

1. Main definition

Exercise 1.1 True or false.

(a) Every linea operator in an n-dimensional vector space has n distinct eigenvalues;

(b) If a matrix has on eigenvalue, it has infinitely many eigenvectors;

(c) There exists a square real matrix with no real eigenvalues;

(d) There exists a square matrix with no (complex) eigenvectors;

(e) Similar matrices always have the same eigenvalues;

(f) Similar matrices always have the same eigenvectors;

(g) The sum of two eigenvectors of a matrix A is always an eigenvector;

(h) The sum of two eigenvectors of a matrix A corresponding to the same eigenvalue λ is always an eigenvector.

Solution. (a) **false**

(b) **true**. Suppose $\mathbf{v}$ be an eigenvector corresponding to the eigenvalue λ of a matrix A, then $A\mathbf{v} = \lambda\mathbf{v}$. For any $\alpha \in \mathbb{F}$, we have

$$A(\alpha\mathbf{v}) = \alpha A\mathbf{v} = \alpha\lambda\mathbf{v} = \lambda(\alpha\mathbf{v})$$

hence $\alpha\mathbf{v}$ is also an eigenvector of λ. Because of the choise of α is abitrary, we conclude that it has infinitely many eigenvectors.

(c) **true**. It's easy to see that the rotation matrix R_γ where $\gamma \neq n\pi$ has no real eigenvalue. Below, we present the special case when $\gamma = \pi/2$. Namely, the rotation matrix about 90°

$$A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The characteristic polynomial

$$\det(A - \lambda I) = (-\lambda)(-\lambda) + 1 = \lambda^2 + 1$$

has roots $\lambda = \pm i$, both roots are complex numbers.

(d) **false.** From fundamental theorem of Algebra, any polynomial with degrees greater than 2 always has complex roots. So, there's always an eigenvalue in complex space.

(e) **true.** As proved in the text book, if matrices A and B are similar, then the determinant

$$\det(A - \lambda I) = \det(B - \lambda I)$$

therefore, A and B have the same eigenvalue.

(f) **true and false.** Since $A \sim B$, they're basically the same transformation but in different basis. We proved earlier that they have the same eigenvalues, therefore the asbtract eigenvector corresponding to this eigenvalue is the same, but numerically different because of the basis we denote them.

(g) **false.** There are plenty of counterexamples, the simplest one is

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

having eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$ with the corresponding eigenvector

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

however their sum $\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is not an eigenvector.

(h) **true.** Suppose that $\mathbf{v}, \mathbf{w}$ are eigenvectors corresponding to an eigenvector λ. Then

$$\begin{aligned} & A\mathbf{v} = \lambda \mathbf{v} \quad \text{and} \quad A\mathbf{w} = \lambda \mathbf{w} \\ \implies \quad & A(\mathbf{v} + \mathbf{w}) = \lambda(\mathbf{v} + \mathbf{w}) \end{aligned}$$

hence $\mathbf{v} + \mathbf{w}$ is also an eigenvector of λ.

□

Exercise 1.5 Prove that eigenvalues (counting multiplicities) of a triangular matrix coincide with its diagonal entries.

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the diagonal entries of the triangular matrix A. Suppose that A is upper triangular matrix,

$$A = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_4 \end{pmatrix}$$

The characteristic polynomial of A is

$$\det(A - \lambda I) = \prod_{i=1}^{n} (\lambda_i - \lambda)$$

and the roots of this polynomial are exactly $\lambda_1, \lambda_2, \ldots, \lambda_n$. For lower triangular matrix, we carry out the same proof. $\square$

Exercise 1.6 An operator A is called *nilpoten* if $A^k = \mathbf{0}$ for some k. Prove that if A is nilpoten, then $\sigma(A) = \{0\}$.

Proof. First, we prove that 0 is an eigenvalue of A, that is $\sigma(A) \neq \varnothing$. Then we show that if $\lambda \in \sigma(A)$, then $\lambda = 0$.

- Using the property of determinant, we have

  $$0 = \det(A^k) = (\det A)^k$$

  hence $\det A = \det(A - 0I) = 0$. We conclude that $0 \in \sigma(A)$.

- Lastly, let $\lambda \in \sigma(A)$ be an abitrary eigenvalue of A. Then there exists $\mathbf{v} \neq \mathbf{0}$ such that

  $$\begin{aligned} & A\mathbf{v} = \lambda \mathbf{v} \\ \implies \quad & A^2\mathbf{v} = A\lambda\mathbf{v} = \lambda A\mathbf{v} = \lambda^2\mathbf{v} \end{aligned}$$

  keep appying this for the total of k times, we'd get $\lambda^k\mathbf{v} = A^k\mathbf{v} = \mathbf{0}$. Since $\mathbf{v}$ is a non zero vector, we must have $\lambda = 0$.

$\square$

Exercise 1.7 Show that the characteristic polynomial of a block triangular matrix

$$M := \begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix}$$

where A and B are square matrices, coincides with $\det(A - \lambda I)\det(B - \lambda I)$.

Proof. Let $A' := A - \lambda I$ and $B' := B - \lambda I$. Therefore, the characteristic polynomial of M is

$$\begin{aligned} \det(M - \lambda I) &= \det \begin{pmatrix} A' & * \\ \mathbf{0} & B' \end{pmatrix} \\ &= \det(A')\det(B') \\ &= \det(A - \lambda I)\det(B - \lambda I) \end{aligned}$$

where the second equality comes from the determinant of block the matrix. $\square$

Exercise 1.8 Let $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ be a basis in a vector space V. Assume also that the first k vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ of the basis are eigenvectors of an operator A, corresponding to an eigenvalue λ. Show that in this basis the matrix of the operator A has block triangular form

$$\begin{pmatrix} \lambda I_k & * \\ \mathbf{0} & B \end{pmatrix}$$

where I_k is $k \times k$ identity matrix and B is some $(n-k) \times (n-k)$ matrix.

Proof. Let $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$. To write A in this basis, we need to know what happens to each basis vector. Now, for the first k vectors, for each $j = 1, 2, \ldots, k$ we have

$$A\mathbf{v_j} = \lambda \mathbf{v_j} = \begin{pmatrix} 0 \\ \vdots \\ \lambda \\ \vdots \\ 0 \end{pmatrix}_{\mathcal{B}}$$

where λ is at the j-th position. And for the others $(n-k)$ vectors, nah we don't really care about them. The matrix A written in basis $\mathcal{B}$ would look like

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & \\ 0 & \lambda_2 & \cdots & 0 & \\ \vdots & \vdots & \ddots & \vdots & \boxed{\text{what ev}} \\ 0 & 0 & \cdots & \lambda_n & \\ & & & & \\ & \boxed{\text{zeros}} & & & \boxed{(n-1)\text{ block}} \end{pmatrix} = \begin{pmatrix} \lambda I_k & * \\ \mathbf{0} & B \end{pmatrix}$$

□

Exercise 1.9 Use the two previous exercises to prove that geometric multiplicity of an eigenvalue cannot exceed its algebraic multiplicity.

Proof. Let $\lambda \in \sigma(A)$ be an arbitrary basis in operator $A : V \to V$. Suppose that $\mathbf{v_1}, \ldots, \mathbf{v_k}$ be basis in subspace $\ker(A - \lambda I)$. Then we can complete this system to a basis in V, hence we can let

$$\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_k}, \ldots, \mathbf{v_n}\}$$

be a basis in V (we implicitely define $\dim V = n$). From the previous exercise, we can write A in that basis $\mathcal{B}$ as follows

$$[A]_{\mathcal{B}} = \begin{pmatrix} \lambda I_k & * \\ \mathbf{0} & B \end{pmatrix}$$

because eigenvalues don't rely on the choise of basis, λ is also the root of characteristic polynomial (with variable x)

$$\begin{aligned} \det(A - xI) &= \det(\lambda I_k - xI_k) \cdot \det(B - xI_{n-k}) \\ &= (\lambda - x)^k \cdot \det(B - xI) \end{aligned}$$

therefore

$$\boxed{\text{algebraic mult.} \geq k = \text{geometric mult.}}$$

□

Exercise 1.10 Prove that determinant of a matrix A is equals to the product of its eigenvalues (counting multiplicities).

Proof. Let this square matrix $A = [a]_{jk}$ and $A' := A - \lambda I = [b]_{jk}$ where $b_{jj} = a_{jj} - \lambda$. As in previous chapter, we defined $\operatorname{perm}(n)$ be to the set of all permutations of n letters. Let e be the indentity permutation in $\operatorname{perm} n$, that is $e(j) = j$ for all j, and $\operatorname{sign} e = 1$. Therefore,

$$\begin{aligned} \det(A - \lambda I) &= \sum_{\sigma \in \operatorname{perm} n} \operatorname{sign} \sigma \cdot \prod_{j=1}^{n} b_{\sigma(j)j} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) + q(\lambda) \qquad (4.1) \end{aligned}$$

where $q(\lambda) := \sum_{\sigma \neq e} \operatorname{sign} \sigma \cdot \prod_{i=1}^{n} a_{\sigma(j)j}$. For any $\sigma \neq e$, there must be at least two letters were swapped. Since all the variables λ's are all on the main diagonal, the polynomial $\prod_{i=1}^{n} a_{\sigma(j)j}$ must have at most $n-2$ degrees. Note also that the coefficient of λ^n in this polynomial is $(-1)^n$.

Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be all the eigenvalues of A, hence they are the roots of $\det(A - \lambda I)$. So

$$\det(A - \lambda I) = \operatorname{coef}(\lambda^n) \cdot (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) \qquad (4.2)$$

Because $\operatorname{coef}(\lambda^n) = (-1)^n$, plugging $\lambda = 0$ in equation (**??**) we'd get

$$\boxed{\det A = \lambda_1 \lambda_2 \cdots \lambda_n}$$

□

Exercise 1.11 Prove that trace of a matrix equals to the sum of its eigenvalue.

Proof. From equation (**??**) in above exercise, we have

$$\det(A - \lambda I) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

has $\lambda_1, \ldots, \lambda_n$ as root. So by Vietta theorem, we obtain that the coefficient of λ^{n-1} is

$$\operatorname{coef}(\lambda^{n-1}) = -\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_n}{(-1)^n}.$$

Now, finding the coefficient of λ^{n-1} in equation (**??**) is similar to find the coefficient of λ^{n-1} in the polynomial

$$p(\lambda) := (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda),$$

and since this polynomial has roots $a_{11}, a_{22}, \ldots, a_{nn}$, using Vietta theorem we'd get

$$\operatorname{coef} \lambda^{n-1} = -\frac{a_{11} + a_{22} + \cdots + a_{nn}}{(-1)^n}.$$

Therefore, we conclude that

$$\boxed{\operatorname{trace}(A) = \sum a_{jj} = \sum \lambda_j}$$

□

2. Diagonalization

Exercise 2.1 Let A be $n \times n$ matrix. True or false.

(a) A^{T} has the same eigenvalues as A;

(b) A^{T} has the same eigenvectors as A;

(c) If A is diagonalizable, the so is A^{T}

Proof. (a) **true.** The characteristic polynomial of A^{T} is

$$\det(A^{\mathrm{T}} - \lambda I) = \det(A - \lambda I)^{\mathrm{T}} = \det(A - \lambda I)$$

hene A and A^{T} have the same eigenvalues.

(b) **false.** Not generally true. In fact, we can check this using the example in the text book (page 112). However, if A is a symmetric matrix ($A^{\mathrm{T}} = A$), then A and A^{T} have the same eigenvectors.

(c) **true.** Since A and A^{T} have the same characteristic polynomial, hence if λ is a root, then $\mathrm{alg}_A(\lambda) = \mathrm{alg}_{A^{\mathrm{T}}}(\lambda)$, i.e. they have the same number of repeated roots. Because A is diagonalizable, it's only the case that

$$\mathrm{alg}_A(\lambda) = \mathrm{geo}_A(\lambda)$$

now it only remains to show that $\mathrm{geo}_A(\lambda) = \mathrm{geo}_{A^{\mathrm{T}}}(\lambda)$. Observe that

$$\dim\ker(A - \lambda I) = n - \mathrm{rank}(A - \lambda I)$$

and

$$\dim\ker(A^{\mathrm{T}} - \lambda I) = n - \mathrm{rank}(A^{\mathrm{T}} - \lambda I)$$

note also that

$$\mathrm{rank}(A^{\mathrm{T}} - \lambda I) = \mathrm{rank}(A - \lambda I)^{\mathrm{T}} = \mathrm{rank}(A - \lambda I)$$

this shows that $\mathrm{geo}_A(\lambda) = \mathrm{geo}_{A^{\mathrm{T}}}(\lambda)$. Therefore, A^{T} is also diagonalizable.

$\square$

Exercise 2.2 Let A be a square matrix with real entries, and let λ be its complex eigenvalue. Suppose that $\mathbf{v} = (v_1, v_2, \ldots, v_n)^{\mathrm{T}}$ is a corresponding eigenvector. Prove that the $\overline{\lambda}$ is an eigenvalue of A and that $A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$.

Proof. Let us separate the real and imaginary part of $\mathbf{v}$ as $\mathbf{v} = \mathbf{a} + i\mathbf{b}$, where $\mathbf{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$ and $\mathbf{b} = (b_1, \ldots, b_n)^{\mathrm{T}}$ are real-valued vectors in $\mathbb{R}^n$. Similarly, we also write λ as $\lambda = \alpha + i\beta$ where $\alpha, \beta \in \mathbb{R}$. Because A is real and λ is its eigenvalue, we obtain that

$$\begin{aligned}
& A\mathbf{v} = \lambda\mathbf{v} \\
\implies \quad & A\mathbf{a} + iA\mathbf{b} = (\alpha + i\beta)(\mathbf{a} + i\mathbf{b}) \\
& \qquad\qquad = (\alpha\mathbf{a} - \beta\mathbf{b}) + i(\beta\mathbf{a} + \alpha\mathbf{b})
\end{aligned}$$

Applying complex conjugate on both sides, we get

$$\begin{aligned} A\mathbf{a} - iA\mathbf{b} &= (\alpha\mathbf{a} - \beta\mathbf{b}) - i(\beta\mathbf{a} + \alpha\mathbf{b}) \\ \implies \quad A\overline{\mathbf{v}} &= (\alpha - i\beta)(\mathbf{a} - i\mathbf{b}) \\ &= \overline{\lambda}\overline{\mathbf{v}} \end{aligned}$$

□

Chapter 5

Inner Product

1. Inner product

2. Orthogonality

3. Orthogonal projection

Exercise 3.12 Show that for a subspace E we have $(E^{\perp})^{\perp} = E$.

Proof. First note that for any $\mathbf{e} \in E$, $\mathbf{e} \perp E^{\perp}$, hence $E \perp E^{\perp}$. To prove this assertion, we'll prove that each set is a subset of one another.

- Let $\mathbf{x} \in E$. From above reasoning we have $\mathbf{x} \perp E^{\perp}$, but this immediately implies that $\mathbf{x} \in (E^{\perp})^{\perp}$ because this set contains all of those vectors that orthogonal to $E^{\perp}$. Thus, $E \subseteq (E^{\perp})^{\perp}$.
- Now suppose that $\mathbf{y} \in (E^{\perp})^{\perp}$. We can decompose $\mathbf{y}$ as the sum

  $$\mathbf{y} = \text{proj}_E \mathbf{y} + \mathbf{y_0}$$

  where $\mathbf{y_0} \perp E$. Note the inner product $\langle \mathbf{y_0}, \mathbf{y_0} \rangle = \langle \mathbf{y}, \mathbf{y_0} \rangle - \langle \text{proj}_E \mathbf{y}, \mathbf{y_0} \rangle$. Since $\mathbf{y} \perp E^{\perp}$, $\mathbf{y_0} \in E^{\perp}$ their inner product $\langle \mathbf{y}, \mathbf{y_0} \rangle = 0$. Similarly since $\text{proj}_E \mathbf{y} \in E$, $\mathbf{y_0} \perp E$, their inner product is also 0. Hence we obtain that

  $$\langle \mathbf{y_0}, \mathbf{y_0} \rangle = 0 \iff \mathbf{y_0} = \mathbf{0}.$$

  This implies that $\mathbf{y} = \text{proj}_E \mathbf{y} \in E$ for all $\mathbf{y}$. Hence $(E^{\perp})^{\perp} \subseteq E$.

$\square$

4. Least square

Exercise 4.5 (*Minimal norm solution*) Let an equation $A\mathbf{x} = \mathbf{b}$ has a solution, and let A has non-trivial kernel (so the solution is not unique.) Prove that

(a) There exist a unique solution $\mathbf{x_0}$ of $A\mathbf{x} = \mathbf{b}$ minimizing the norm $\|\mathbf{x}\|$, i.e. that there exists unique $\mathbf{x_0}$ such that $A\mathbf{x_0} = \mathbf{b}$ and $\|\mathbf{x_0}\| \le \|\mathbf{x}\|$ for any $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{b}$.

(b) $\mathbf{x_0} = \text{proj}_{(\ker A)^\perp}\mathbf{x}$ for any $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{b}$.

Proof.

(a) Let V be the domain space of A, and let $\mathcal{S} \subset V$ be the set of all solution of $A\mathbf{x} = \mathbf{b}$. Let arbitrary $\mathbf{w} \in \mathcal{S}$, and denote $\mathbf{x_0} \in (\ker A)^\perp$ be its projection onto subspace $(\ker A)^\perp$. By decomposition, we obtain that there is $\mathbf{x_h} \in ((\ker A)^\perp)^\perp = \ker A$, such that

$$\begin{aligned} & \mathbf{w} = \mathbf{x_0} + \mathbf{x_h} \\ \implies \quad & A\mathbf{x_0} = A\mathbf{w} - A\mathbf{x_h} = \mathbf{b} - \mathbf{0} = \mathbf{b} \end{aligned}$$

So $\mathbf{x_0} \in (\ker A)^\perp$ is also a solution of the above equation ($\mathbf{x_0}$ is a more special root than our regular $\mathbf{w}$).

In words, there exists a solution $\mathbf{x_0}$ such that $\mathbf{x_0} \in (\ker A)^\perp$. If $\mathbf{x_0} = \mathbf{0}$, the above inequality is true. So from now on, we only consider when $\mathbf{x_0} \ne \mathbf{0}$. Note that, for any general solution $\mathbf{x} \in \mathcal{S}$ can be written as $\mathbf{x} = \mathbf{x_0} + \mathbf{e}$ for some vector $\mathbf{e} \in \ker A$ (proved in chapter 2). Then

$$\begin{aligned} \|\mathbf{x_0}\|^2 &= \langle \mathbf{x_0}, \mathbf{x_0} \rangle = \langle \mathbf{x} - \mathbf{e},\ \mathbf{x_0} \rangle \\ &= \langle \mathbf{x}, \mathbf{x_0} \rangle - \langle \mathbf{e}, \mathbf{x_0} \rangle \\ &= \langle \mathbf{x}, \mathbf{x_0} \rangle && \text{(because } \mathbf{x_0} \perp \mathbf{e}\text{)} \\ &\le \|\mathbf{x}\| \cdot \|\mathbf{x_0}\| && \text{(using Cauchy–Schwarz)} \end{aligned}$$

divided both sides by $\|\mathbf{x_0}\| \ne 0$, we get the desired inequality $\|\mathbf{x_0}\| \le \|\mathbf{x}\|$ for any $\mathbf{x} \in \mathcal{S}$.

(b) Let $\mathbf{w_1}, \mathbf{w_2} \in \mathcal{S}$ be solutions of $A\mathbf{x} = \mathbf{b}$. We wish to show that their projection $P\mathbf{w_1}$ and $P\mathbf{w_2}$ onto $(\ker A)^\perp$ are the same. Because $(\ker A)^\perp$ is itself a subspace, then

$$P\mathbf{w_1} - P\mathbf{w_2} \in (\ker A)^\perp. \tag{5.1}$$

Because $P\mathbf{w_1}, P\mathbf{w_2}$ are projections, we obtain that there exists $\mathbf{e_1}, \mathbf{e_2} \in ((\ker A)^\perp)^\perp = \ker A$, such that

$$\mathbf{w_1} = P\mathbf{w_1} + \mathbf{e_1}$$

and

$$\mathbf{w_2} = P\mathbf{w_2} + \mathbf{e_2}.$$

Substracting these two equations, and applying A from both sides, we get

$$\begin{aligned} A(P\mathbf{w_1} - P\mathbf{w_2}) &= A\mathbf{w_1} - A\mathbf{w_2} - A\mathbf{e_1} + A\mathbf{e_2} \\ &= \mathbf{b} - \mathbf{b} - \mathbf{0} + \mathbf{0} = \mathbf{0}. \end{aligned}$$

This implies that

$$P\mathbf{w_1} - P\mathbf{w_2} \in \ker A. \tag{5.2}$$

But $\ker A \oplus (\ker A)^\perp = V$, hence $(\ker A) \cap (\ker A)^\perp = \{\mathbf{0}\}$ therefore

$$P\mathbf{w_1} - P\mathbf{w_2} \in (\ker A) \cap (\ker A)^\perp = \{0\}$$

and thus $P\mathbf{w_1} = P\mathbf{w_2}$ as expected.

□