Chapter 6: Vector Space and Subspace

6.1 Vectors

Let $\mathbb{R}^m = \left\{ \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \middle| \; v_i \in \mathbb{R}, 1 \le i \le m \right\}$ is the set of all column vectors of length m with entries from $\mathbb{R}$.
$\mathbb{R}^m$ is also called the *Euclidean m-space*. This is the same set $M_{m,1}(\mathbb{R})$.

For convenience, we often identify the set $\mathbb{R}^m$ with the Cartesian product of m identical real number set, i.e., $\underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{m \text{ times}} = \{(v_1, \ldots, v_m) \mid v_i \in \mathbb{R}, 1 \le i \le m\}$.

Theorem 6.1.1: *Under the addition and scalar multiplication of matrix, we have*

1. *Additive Closure: If $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{R}^m$, then $\boldsymbol{u} + \boldsymbol{v} \in \mathbb{R}^m$.*
2. *Commutativity: If $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{R}^m$, then $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$.*
3. *Additive Associativity: If $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w} \in \mathbb{R}^m$, then $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$.*
4. *Zero Vector: There is a vector, $\mathbf{0}_m$, called the* zero vector*, such that $\boldsymbol{u} + \mathbf{0}_m = \boldsymbol{u}$ for all $\boldsymbol{u} \in \mathbb{R}^m$.*
5. *Additive Inverses: If $\boldsymbol{u} \in \mathbb{R}^m$, then there exists a vector $-\boldsymbol{u} \in \mathbb{R}^m$ so that $\boldsymbol{u} + (-\boldsymbol{u}) = \mathbf{0}_m$.*
6. *Scalar Closure: If $a \in \mathbb{R}$ and $\boldsymbol{u} \in \mathbb{R}^m$, then $a\boldsymbol{u} \in \mathbb{R}^m$.*
7. *Scalar Multiplication Associativity: If $a, b \in \mathbb{R}$ and $\boldsymbol{u} \in \mathbb{R}^m$, then $a(b\boldsymbol{u}) = (ab)\boldsymbol{u}$.*
8. *Distributivity across Vector Addition: If $a \in \mathbb{R}$ and $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{R}^m$, then $a(\boldsymbol{u} + \boldsymbol{v}) = a\boldsymbol{u} + a\boldsymbol{v}$.*
9. *Distributivity across Scalar Addition: If $a, b \in \mathbb{R}$ and $\boldsymbol{u} \in \mathbb{R}^m$, then $(a + b)\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{u}$.*
10. *One: If $\boldsymbol{u} \in \mathbb{R}^m$, then $1\boldsymbol{u} = \boldsymbol{u}$.*

Definition 6.1.2: A *vector space V over* $\mathbb{R}$ is a non-empty set with two laws of combination called *vector addition* "+" (or simply addition) and *scalar multiplication* "·" satisfying the following axioms:

(V1) $+ : V \times V \to V$ is a mapping and $+(\alpha, \beta)$ written by $\alpha + \beta$ is called the *sum* of α and β.

(V2) $+$ is associative.

(V3) $+$ is commutative.

(V4) There is an element, denoted by $\mathbf{0}$, such that $\alpha + \mathbf{0} = \alpha$ for all $\alpha \in V$. Note that such vector is unique. It is called the *zero vector* of V.

(V5) For each $\alpha \in V$ there is an element in V, denoted by $-\alpha$ such that $\alpha + (-\alpha) = \mathbf{0}$.

(V6) $\cdot : \mathbb{R} \times V \to V$ is a mapping which associates $a \in \mathbb{R}$ and $\alpha \in V$ a unique element denoted by $a \cdot \alpha$ or simply $a\alpha$ in V. This mapping is called the *scalar multiplication.*

(V7) Scalar multiplication is associative, i.e., $a(b\alpha) = (ab)\alpha$ for all $a, b \in \mathbb{R}$, $\alpha \in V$.

(V8) Scalar multiplication is distributive with respect to $+$, i.e., $a(\alpha+\beta) = a\alpha + a\beta$ for all $a \in \mathbb{R}$, $\alpha, \beta \in V$.

(V9) For each $a, b \in \mathbb{R}$, $\alpha \in V$, $(a+b)\alpha = a\alpha + b\alpha$.

(V10) For each $\alpha \in V$, $1 \cdot \alpha = \alpha$.

Elements of V and $\mathbb{R}$ are called *vectors* and *scalars*, respectively. After this section, vectors are often denoted by lower case Greek letters $\alpha, \beta, \gamma, \ldots$ (you may still use $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \ldots$) and scalars are often denoted by lower case Latin letters $a, b, c, \ldots$.

Lemma 6.1.3 (Cancellation Law)**:** *Suppose α, β and γ are vectors in a vector space. If $\alpha+\beta = \alpha+\gamma$, then $\beta = \gamma$.*

Corollary 6.1.4: *Suppose α and β are vectors in a vector space. If $\alpha + \beta = \alpha$, then $\beta = \mathbf{0}$.*

Proposition 6.1.5: *Let V be a vector space over $\mathbb{R}$. We have*

(a) $\forall \alpha \in V$, $0\alpha = \mathbf{0}$.

(b) $\forall \alpha \in V$, $(-1)\alpha = -\alpha$.

(c) $\forall a \in \mathbb{R}$, $a\mathbf{0} = \mathbf{0}$.

Examples:

1. The set $\{0\}$ is a vector space over $\mathbb{R}$.

2. Let n be a positive integer. $\mathbb{R}^n$ is a vector space over $\mathbb{R}$. In particular, $\mathbb{R}$ is a vector space over $\mathbb{R}$.

3. Let m and n be positive integers. The set $M_{m,n}(\mathbb{R})$ is a vector space over $\mathbb{R}$ under the usual addition and scalar multiplication.

4. Suppose $\mathscr{I}$ is an interval of $\mathbb{R}$. Let $C^0(\mathscr{I})$ be the set of all continuous real valued functions defined on $\mathscr{I}$. Then $C^0(\mathscr{I})$ is a vector space over $\mathbb{R}$.

5. Let $\mathbb{R}[x]$ be the set of all polynomials in the indeterminate x over $\mathbb{R}$. Under the usual addition and scalar multiplication of polynomials, $\mathbb{R}[x]$ is a vector space over $\mathbb{R}$.

6. Let n be a positive integer. Let P_n be the subset of $\mathbb{R}[x]$ consisting of all polynomials in x of degree n or less (of course, together with the zero polynomial). Then P_n is a vector space over $\mathbb{R}$ with the same addition and scalar multiplication as in $\mathbb{R}[x]$ defined in the previous example. Namely, P_n can be written as $\left\{ \sum_{i=0}^{n} a_i x^i \;\middle|\; a_i \in \mathbb{R} \right\}$.

7. **(MATH Major)** I am free to define my set and my operations any way I please. They may not look natural, or even useful, but we will now verify that they provide us with another example of a vector space. We will check all it satisfies all the definition of vector spaces.

   (The crazy vector space) Let $C = \{(x_1, x_2) \mid x_1,\, x_2 \in \mathbb{R}\}$.

   (a) Vector Addition: $(x_1,\, x_2) \oplus (y_1,\, y_2) = (x_1 + y_1 + 1,\, x_2 + y_2 + 1)$.

   (b) Scalar Multiplication: $a \cdot (x_1,\, x_2) = (ax_1 + a - 1,\, ax_2 + a - 1)$, where $a \in \mathbb{R}$.

   - Property (V1) and (V6):
     The result of each operation is a pair of real numbers, so these two closure properties are fulfilled.

   - Property (V2):

$$
\begin{aligned}
\boldsymbol{u} \oplus (\boldsymbol{v} \oplus \boldsymbol{w}) &= (x_1,\, x_2) \oplus ((y_1,\, y_2) \oplus (z_1,\, z_2)) \\
&= (x_1,\, x_2) \oplus (y_1 + z_1 + 1,\, y_2 + z_2 + 1) \\
&= (x_1 + (y_1 + z_1 + 1) + 1,\, x_2 + (y_2 + z_2 + 1) + 1) \\
&= (x_1 + y_1 + z_1 + 2,\, x_2 + y_2 + z_2 + 2) \\
&= ((x_1 + y_1 + 1) + z_1 + 1,\, (x_2 + y_2 + 1) + z_2 + 1) \\
&= (x_1 + y_1 + 1,\, x_2 + y_2 + 1) \oplus (z_1,\, z_2) \\
&= ((x_1,\, x_2) \oplus (y_1,\, y_2)) \oplus (z_1,\, z_2) \\
&= (\boldsymbol{u} \oplus \boldsymbol{v}) \oplus \boldsymbol{w}
\end{aligned}
$$

   - Property (V3):

$$
\begin{aligned}
\boldsymbol{u} \oplus \boldsymbol{v} &= (x_1,\, x_2) \oplus (y_1,\, y_2) = (x_1 + y_1 + 1,\, x_2 + y_2 + 1) \\
&= (y_1 + x_1 + 1,\, y_2 + x_2 + 1) = (y_1,\, y_2) \oplus (x_1,\, x_2) \\
&= \boldsymbol{v} \oplus \boldsymbol{u}
\end{aligned}
$$

   - Property (V4):

     The zero vector is $\mathbf{0} = (-1, -1)$ (**not** $(0, 0)$)

$$
\boldsymbol{u} \oplus \mathbf{0} = (x_1,\, x_2) \oplus (-1,\, -1) = (x_1 + (-1) + 1,\, x_2 + (-1) + 1) = (x_1,\, x_2) = \boldsymbol{u}
$$

- Property (V5):
  For each vector, $\boldsymbol{u}$, we must locate an additive inverse, $\boldsymbol{-u}$. Here it is, $-(x_1, x_2) = (-x_1 - 2, -x_2 - 2)$. As odd as it may look, I hope you are withholding judgment. Check:

$$\begin{aligned}\boldsymbol{u} \oplus (\boldsymbol{-u}) &= (x_1, x_2) \oplus (-x_1 - 2, -x_2 - 2)\\ &= (x_1 + (-x_1 - 2) + 1, x_2 + (-x_2 - 2) + 1) = (-1, -1) = \mathbf{0}\end{aligned}$$

- Property (V7):

$$\begin{aligned}a \cdot (b \cdot \boldsymbol{u}) &= a \cdot (b \cdot (x_1, x_2))\\ &= a \cdot (bx_1 + b - 1, bx_2 + b - 1)\\ &= (a(bx_1 + b - 1) + a - 1, a(bx_2 + b - 1) + a - 1)\\ &= ((abx_1 + ab - a) + a - 1, (abx_2 + ab - a) + a - 1)\\ &= (abx_1 + ab - 1, abx_2 + ab - 1)\\ &= (ab) \cdot (x_1, x_2)\\ &= (ab) \cdot \boldsymbol{u}\end{aligned}$$

- Property (V8):

$$\begin{aligned}&a \cdot (\boldsymbol{u} \oplus \boldsymbol{v})\\ &\quad = a \cdot ((x_1, x_2) \oplus (y_1, y_2))\\ &\quad = a \cdot (x_1 + y_1 + 1, x_2 + y_2 + 1)\\ &\quad = (a(x_1 + y_1 + 1) + a - 1, a(x_2 + y_2 + 1) + a - 1)\\ &\quad = (ax_1 + ay_1 + a + a - 1, ax_2 + ay_2 + a + a - 1)\\ &\quad = (ax_1 + a - 1 + ay_1 + a - 1 + 1, ax_2 + a - 1 + ay_2 + a - 1 + 1)\\ &\quad = ((ax_1 + a - 1) + (ay_1 + a - 1) + 1, (ax_2 + a - 1) + (ay_2 + a - 1) + 1)\\ &\quad = (ax_1 + a - 1, ax_2 + a - 1) \oplus (ay_1 + a - 1, ay_2 + a - 1)\\ &\quad = a \cdot (x_1, x_2) \oplus a \cdot (y_1, y_2)\\ &\quad = a \cdot \boldsymbol{u} \oplus a \cdot \boldsymbol{v}\end{aligned}$$

- Property (V9):

$$\begin{aligned}&(a + b) \cdot \boldsymbol{u}\\ &\quad = (a + b) \cdot (x_1, x_2)\\ &\quad = ((a + b)x_1 + (a + b) - 1, (a + b)x_2 + (a + b) - 1)\\ &\quad = (ax_1 + bx_1 + a + b - 1, ax_2 + bx_2 + a + b - 1)\\ &\quad = (ax_1 + a - 1 + bx_1 + b - 1 + 1, ax_2 + a - 1 + bx_2 + b - 1 + 1)\\ &\quad = ((ax_1 + a - 1) + (bx_1 + b - 1) + 1, (ax_2 + a - 1) + (bx_2 + b - 1) + 1)\\ &\quad = (ax_1 + a - 1, ax_2 + a - 1) \oplus (bx_1 + b - 1, bx_2 + b - 1)\\ &\quad = a \cdot (x_1, x_2) \oplus b \cdot (x_1, x_2)\\ &\quad = a \cdot \boldsymbol{u} \oplus b \cdot \boldsymbol{u}\end{aligned}$$

- Property (V10):

$$1 \cdot \boldsymbol{u} = 1 \cdot (x_1, x_2) = (x_1 + 1 - 1, x_2 + 1 - 1) = (x_1, x_2) = \boldsymbol{u}$$

Thus, C is a vector space, as crazy as that may seem.

Notice that in the case of the zero vector and additive inverses, we only had to propose possibilities and then verify that they were the correct choices. You might try to discover how you would arrive at these choices, though you should understand why the process of discovering them is not a necessary component of the proof itself.

6.2 Subspaces

Definition 6.2.1: Let V be vector space. A subset W of V is said to be a *subspace* of V if

1. W is nonempty.
2. For $\alpha, \beta \in W$, $\alpha + \beta \in W$.
3. For $a \in \mathbb{R}$ and $\alpha \in W$, $a\alpha \in W$.

Proposition 6.2.2: *Let V be a vector space and W a subspace of V. Then $\mathbf{0}$ is in W.*

Proof: By Definition 6.2.1 Condition 1, W is nonempty. So we may choose $\alpha \in W$.
By Definition 6.2.1 Condition 3, with $a = -1$, $(-1)\alpha = -\alpha \in W$.
By Definition 6.2.1 Condition 2, choose $\beta = -\alpha$. Then $\alpha + \beta \in W$. But $\alpha + \beta = \alpha + (-\alpha) = \mathbf{0}$. So $\mathbf{0} \in W$. □

Theorem 6.2.3: *Let V be a vector space and W a subset of V. W is a subspace if and only if*

(a) W is nonempty.

(b) For any $a \in \mathbb{R}$, $\alpha, \beta \in W$, $a\alpha + \beta \in W$.

Example 6.2.1: $V = \mathbb{R}^m$, $W = \{\mathbf{0}_m\}$.
W consists of one element. It is called the *zero subspace* of V.
Check that it is a subspace: W is nonempty. For any $a \in \mathbb{R}$, $\alpha, \beta \in W$, $\alpha = \beta = \mathbf{0}_m$, $a\alpha + \beta = \mathbf{0}_m \in W$. Thus by Theorem 6.2.3, W is a subspace. ■

Example 6.2.2: $V = \mathbb{R}^m$, $W = V$. Obviously W is a subspace. ■

Example 6.2.3: $V = \mathbb{R}^m$, $W = \{\alpha \in V \mid [\alpha]_1 = 0\}$.

Check that W is a subspace:

Example 6.2.4: We identify $\mathbb{R}^3$ with $\mathbb{R}\times\mathbb{R}\times\mathbb{R}$. Let $V = \mathbb{R}^3$ and $W = \{(x, y, z) \in \mathbb{R}^3 \mid x+2y+3z = 0\}$.

Example 6.2.5: $V = \mathbb{R}^m$, $W = \{\alpha \in V \mid [\alpha]_1 = 1\}$.

Example 6.2.6: $V = \mathbb{R}^m$, $W = \{(v_1, \ldots, v_m)^t \in V \mid \sum_{i=1}^n v_i = 1\}$.

Theorem 6.2.4: *Let $A \in M_{m,n}$. $\mathcal{N}(A)$ is a subspace of $\mathbb{R}^n$.*

Example 6.1 Math major only: Let S_n be the set of symmetric matrices of $M_{n,n}$. Then S_n is a subspace of $M_{n,n}$.
Checking: Since $O_n \in W$, W is nonempty. Suppose $a \in \mathbb{R}$, $A, B \in W$. Then $A^t = A$, $B^t = B$.

$$(aA + B)^t = aA^t + B^t = aA + B.$$

Thus $aA + B \in S_n$. Hence S_n is a subspace by Theorem 6.2.3.

Example 6.2 Math major only: Let

$$F = \{f(x) \in P_n \mid f(1) = 0\}.$$

Then F is a subspace of P_n.
Checking: Since the zero polynomial is in F, F is nonempty. Suppose $a \in \mathbb{R}$, $f, g \in F$. Then $f(1) = g(1) = 0$. Hence

$$(af + g)(1) = af(1) + g(1) = a0 + 0 = 0.$$

So $af + g \in F$. Hence F is a subspace by Theorem 6.2.3.

Example 6.3 Math major only: Let

$$E = \{f(x) \in P_n \mid f(x) = f(-x)\}.$$

Then E is a subspace of P_n.
Checking: Since $0 \in E$, E is nonempty. Suppose $a \in \mathbb{R}$, $f, g \in E$. Then $f(x) = f(-x)$, $g(x) = g(-x)$. Hence

$$(af + g)(-x) = af(-x) + g(-x) = af(x) + g(x) = (af + g)(x).$$

So $af + g \in E$. Hence E is a subspace by Theorem 6.2.3.

6.3 Linear Combinations

Definition 6.3.1: Given n vectors $\alpha_1, \ldots, \alpha_n \in \mathbb{R}^m$ and n scalars $c_1, \ldots, c_n \in \mathbb{R}$, their *linear combination* is the vector

$$c_1\alpha_1 + \cdots + c_n\alpha_n = \sum_{i=1}^{n} c_i\alpha_i.$$

Example 6.3.1: Suppose that

$$c_1 = 1, c_2 = -4, c_3 = 2, c_4 = -1 \text{ and } \alpha_1 = \begin{bmatrix} 2 \\ 4 \\ -3 \\ 1 \\ 2 \\ 9 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 6 \\ 3 \\ 0 \\ -2 \\ 1 \\ 4 \end{bmatrix}, \alpha_3 = \begin{bmatrix} -5 \\ 2 \\ 1 \\ 1 \\ -3 \\ 0 \end{bmatrix}, \alpha_4 = \begin{bmatrix} 3 \\ 2 \\ -5 \\ 7 \\ 1 \\ 3 \end{bmatrix}.$$

Their linear combination is

$$c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 + c_4\alpha_4 = (1)\begin{bmatrix} 2 \\ 4 \\ -3 \\ 1 \\ 2 \\ 9 \end{bmatrix} + (-4)\begin{bmatrix} 6 \\ 3 \\ 0 \\ -2 \\ 1 \\ 4 \end{bmatrix} + (2)\begin{bmatrix} -5 \\ 2 \\ 1 \\ 1 \\ -3 \\ 0 \end{bmatrix} + (-1)\begin{bmatrix} 3 \\ 2 \\ -5 \\ 7 \\ 1 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 4 \\ -3 \\ 1 \\ 2 \\ 9 \end{bmatrix} + \begin{bmatrix} -24 \\ -12 \\ 0 \\ 8 \\ -4 \\ -16 \end{bmatrix} + \begin{bmatrix} -10 \\ 4 \\ 2 \\ 2 \\ -6 \\ 0 \end{bmatrix} + \begin{bmatrix} -3 \\ -2 \\ 5 \\ -7 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} -35 \\ -6 \\ 4 \\ 4 \\ -9 \\ -10 \end{bmatrix}.$$

■

What vectors were you able to create? Do you think you could create the vector $\beta = (13, 15, 5, -17, 2, 25)^t$ in $\mathbb{R}^6$ with a suitable choice of four scalars? Do you think you could create *any* possible vector from $\mathbb{R}^6$ by choosing the proper scalars? These last two questions are very fundamental, and time spent considering them now will prove beneficial later.

Example 6.3.2: The system of linear equation

$$\begin{array}{rcrcrcr} -7x_1 & - & 6x_2 & - & 12x_3 & = & -33 \\ 5x_1 & + & 5x_2 & + & 7x_3 & = & 24 \\ x_1 & & & + & 4x_3 & = & 5 \end{array}$$

i.e.,

$$\begin{pmatrix} -7x_1 - 6x_2 - 12x_3 \\ 5x_1 + 5x_2 + 7x_3 \\ x_1 + 4x_3 \end{pmatrix} = \begin{pmatrix} -33 \\ 24 \\ 5 \end{pmatrix}.$$

It can be rewritten as

$$\begin{pmatrix} -7x_1 \\ 5x_1 \\ x_1 \end{pmatrix} + \begin{pmatrix} -6x_2 \\ 5x_2 \\ 0x_2 \end{pmatrix} + \begin{pmatrix} -12x_3 \\ 7x_3 \\ 4x_3 \end{pmatrix} = \begin{pmatrix} -33 \\ 24 \\ 5 \end{pmatrix}$$

or

$$x_1\begin{pmatrix} -7 \\ 5 \\ 1 \end{pmatrix} + x_2\begin{pmatrix} -6 \\ 5 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} -12 \\ 7 \\ 4 \end{pmatrix} = \begin{pmatrix} -33 \\ 24 \\ 5 \end{pmatrix}.$$

It is known that the system has only solution

$$x_1 = -3, \quad x_2 = 5, \quad x_3 = 2.$$

So, in the context of this example, we can express the fact that these values of the unknowns are a solution by writing the linear combination,

$$(-3)\begin{pmatrix} -7 \\ 5 \\ 1 \end{pmatrix} + (5)\begin{pmatrix} -6 \\ 5 \\ 0 \end{pmatrix} + (2)\begin{pmatrix} -12 \\ 7 \\ 4 \end{pmatrix} = \begin{pmatrix} -33 \\ 24 \\ 5 \end{pmatrix}.$$

Notice how the three vectors in this example are the columns of the coefficient matrix of the system of equations. This is our first hint of the important interplay between the vectors that form the columns of a matrix, and the matrix itself. ■

Example 6.3.3:

$$\begin{array}{rcrcrcl} x_1 & - & x_2 & + & 2x_3 & = & 1 \\ 2x_1 & + & x_2 & + & x_3 & = & 8 \\ x_1 & + & x_2 & & & = & 5 \end{array}$$

can be written as

$$\begin{pmatrix} x_1 - x_2 + 2x_3 \\ 2x_1 + x_2 + x_3 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ 5 \end{pmatrix}.$$

Converting the left-hand side into a linear combination

$$x_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ 5 \end{pmatrix}.$$

Row-reducing the augmented matrix for the system leads to the conclusion that the system is consistent and has free variables, hence infinitely many solutions. For example, the two solutions

$$x_1 = 2, \quad x_2 = 3, \quad x_3 = 1;$$

$$x_1 = 3, \quad x_2 = 2, \quad x_3 = 0$$

can be used together to say that,

$$(2)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + (3)\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + (1)\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ 5 \end{pmatrix} = (3)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + (2)\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + (0)\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

Ignore the middle of this equation, and move all the terms to the left-hand side,

$$(2)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + (3)\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + (1)\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + (-3)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + (-2)\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + (-0)\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Regrouping gives

$$(-1)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + (1)\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + (1)\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Notice that these three vectors are the columns of the coefficient matrix for the system of equations. This equality says there is a linear combination of those columns that equals the vector of all zeros. Give it some thought, but this says that

$$x_1 = -1, \quad x_2 = 1, \quad x_3 = 1$$

is a nontrivial solution to the homogeneous system of equations with the coefficient matrix for the original system. In particular, this demonstrates that this coefficient matrix is singular.

Theorem 6.3.2: *Let $A \in M_{m,n}(\mathbb{R})$. $\alpha = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$ is a solution to the linear system of equations $\mathcal{LS}(A, \boldsymbol{b})$ if and only if $\boldsymbol{b}$ equals the linear combination of the columns of A formed with the entries of α,*

$$\boldsymbol{b} = x_1 A_{*1} + x_2 A_{*2} + \cdots + x_n A_{*n} = \sum_{i=1}^{n} x_i A_{*i}.$$

Proof: Since $A\alpha = \sum_{i=1}^{n} x_i A_{*i}$, we have the theorem. □

Example 6.3.4: Let

$$\alpha_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \alpha_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \alpha_4 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \beta = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

1. Determine if β is a linear combination of $\{\alpha_1, \alpha_2, \alpha_3\}$. If yes, find the linear combination.
2. Determine if β is a linear combination of $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$. If yes, find the linear combination.

Back to see

Example 3.2.9 Solve the following system of linear equations over $\mathbb{R}$:

$$\begin{cases} x_1 + x_2 - 4x_3 + x_4 = 3 \\ 2x_1 - 3x_2 + 7x_3 + 7x_4 = -4 \\ \quad\quad x_2 - 3x_3 - x_4 = 2 \end{cases}$$

We get

$$\begin{aligned} x_1 &= 1 + x_3 - 2x_4 \\ x_2 &= 2 + 3x_3 + x_4 \end{aligned}$$

and the general solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1+a-2b \\ 2+3a+b \\ a \\ b \end{pmatrix}.$$

In this case, $r = 2$, $D = \{1, 2\}$ and $F = \{3, 4\}$.

Now the general solution can be expressed as a linear combination

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + a \begin{pmatrix} 1 \\ 3 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

We will develop the same linear combination using three steps. While the method above is instructive, the method below will be our preferred approach.

Step 1. Write the vector of unknowns as a fixed vector, plus a linear combination of $n - r$ vectors, using the free variables as the scalars.

$$\boldsymbol{z} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix} + x_3 \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix} + x_4 \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix}$$

Step 2. Use 0's and 1's to ensure equality for the entries of the vectors with indices in F (corresponding to the free variables).

$$\boldsymbol{z} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \\ \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} \\ \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} \\ \\ 0 \\ 1 \end{pmatrix}$$

Step 3. For each lead variable, use the augmented matrix to formulate an equation expressing the lead variable as a constant plus multiples of the free variables. Convert this equation into entries of the vectors that ensure equality for each lead variable, one at a time.

$$x_1 \;=\; 1 \;+\; x_3 \;-\; 2x_4 \;\Rightarrow\; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ \\ 0 \\ 1 \end{pmatrix}$$

$$x_2 \;=\; 2 \;+\; 3x_3 \;+\; x_4 \;\Rightarrow\; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 3 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

It only takes us *three* vectors to describe the entire infinite solution set

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 3 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; x_3, x_4 \in \mathbb{R} \right\}.$$

■

One more example:

Example 6.3.5: Consider a linear system of $m = 5$ equations in $n = 7$ unknowns $x_1, \ldots, x_7$, having the augmented matrix A.

$$(A|\boldsymbol{b}) = \left(\begin{array}{rrrrrrr|r} 2 & 1 & -1 & -2 & 2 & 1 & 5 & 21 \\ 1 & 1 & -3 & 1 & 1 & 1 & 2 & -5 \\ 1 & 2 & -8 & 5 & 1 & 1 & -6 & -15 \\ 3 & 3 & -9 & 3 & 6 & 5 & 2 & -24 \\ -2 & -1 & 1 & 2 & 1 & 1 & -9 & -30 \end{array}\right).$$

$$\text{rref}(A|\boldsymbol{b}) = \left(\begin{array}{rrrrrrr|r} \textcircled{1} & 0 & 2 & -3 & 0 & 0 & 9 & 15 \\ 0 & \textcircled{1} & -5 & 4 & 0 & 0 & -8 & -10 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 0 & -6 & 11 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{1} & 7 & -21 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

So $\text{rank}(A) = r = 4 = \text{rank}(A|\boldsymbol{b})$, also $D = \{1, 2, 5, 6\}$ and $F = \{3, 4, 7, 8\}$. Thus, x_1, x_2, x_5, x_6 are lead variables while x_3, x_4, x_7 are free variables. We try to apply our preferred approach to express the general solution.

Theorem 6.3.3: *Suppose that $(A|\boldsymbol{b})$ is the augmented matrix for a consistent linear system of m equations in n unknowns. Let $(H|\boldsymbol{a}) = \operatorname{rref}(A|\boldsymbol{b})$. Suppose that H has r leading columns, with indices $D = \{k_1, \ldots, k_r\}$, while the $n-r$ non-leading columns have indices in $F = \{f_1, \ldots, f_{n-r}, n+1\}$. Define vectors β, α_j, $1 \le j \le n-r$ of length n by*

$$[\beta]_i = \begin{cases} 0 & \text{if } i \in F \\ [\boldsymbol{a}]_k & \text{if } i \in D,\ i = d_k \end{cases}$$

$$[\alpha_j]_i = \begin{cases} 1 & \text{if } i \in F,\ i = f_j \\ 0 & \text{if } i \in F,\ i \ne f_j \\ -[H]_{k,f_j} & \text{if } i \in D,\ i = d_k \end{cases}.$$

Then the set of solutions to the system of equations $A\boldsymbol{x} = \boldsymbol{b}$ is

$$S = \{\beta + c_1\alpha_1 + \cdots + c_{n-r}\alpha_{n-r} \mid c_1, \ldots, c_{n-r} \in \mathbb{R}\} = \left\{ \beta + \sum_{j=1}^{n-r} c_j\alpha_j \;\middle|\; c_j \in \mathbb{R}, 1 \le i \le n-r \right\}.$$

Proof: We need only show that S is the solution set for the system $H\boldsymbol{x} = \boldsymbol{a}$. Only the first r equations of this system we have to consider.

Note that $[H]_{\ell,d_i} = 0$ except $[H]_{\ell,d_\ell} = 1$.

Now $\forall \gamma \in S$, $\gamma = \beta + \sum\limits_{j=1}^{n-r} c_j\alpha_j$ for some $c_j \in \mathbb{R}$.

Consider the ℓ entry of $H\gamma$, $1 \le \ell \le r$:

$$\begin{aligned} [H\gamma]_\ell &= [H(\beta + \sum_{j=1}^{n-r} c_j\alpha_j)]_\ell = [H\beta]_\ell + \sum_{j=1}^{n-r} c_j[H\alpha_j]_\ell = \sum_{i=1}^{n}[H]_{\ell,i}[\beta]_i + \sum_{j=1}^{n-r} c_j \sum_{i=1}^{n}[H]_{\ell,i}[\alpha_j]_i \\ &= \sum_{k=1}^{r}[H]_{\ell,d_k}[\beta]_{d_k} + \sum_{j=1}^{n-r} c_j \left(\sum_{k=1}^{r}[H]_{\ell,d_k}[-[H]_{k,f_j}] + [H]_{\ell,f_j} \right) \\ &= [\beta]_{d_\ell} + \sum_{j=1}^{n-r} c_j \left([H]_{\ell,d_\ell}[-[H]_{\ell,f_j}] + [H]_{\ell,f_j} \right) = [\boldsymbol{a}]_\ell + \sum_{j=1}^{n-r} c_j \left(-[H]_{\ell,f_j} + [H]_{\ell,f_j} \right) = [\boldsymbol{a}]_\ell. \end{aligned}$$

So γ is a solution of $H\boldsymbol{x} = \boldsymbol{a}$.

For the second part of the proof, we let $\gamma = (x_1, x_2, \ldots, x_n)^t$ be a solution of $H\boldsymbol{x} = \boldsymbol{a}$.

Consider each entry of the matrix equation that γ makes equation ℓ of the system true for all $1 \le \ell \le m$,

$$[H]_{\ell,1}x_1 + [H]_{\ell,2}x_2 + \cdots + [H]_{\ell,n}x_n = \sum_{j=1}^{n}[H]_{\ell,j}x_j = [\boldsymbol{a}]_\ell$$

When $\ell \le r$, the leading columns of H have zero entries in row ℓ with the exception of column d_ℓ, which will contain a 1. So for $1 \le \ell \le r$, equation ℓ simplifies to

$$1x_{d_\ell} + [H]_{\ell,f_1}x_{f_1} + [H]_{\ell,f_2}x_{f_2} + \cdots + [H]_{\ell,f_{n-r}}x_{f_{n-r}} = x_{d_\ell} + \sum_{j=1}^{n-r}[H]_{\ell,f_j}x_{f_j} = [\boldsymbol{a}]_\ell.$$

Thus

$$[\gamma]_{d_\ell} = x_{d_\ell} = [\boldsymbol{a}]_\ell + \sum_{j=1}^{n-r}(-[H]_{\ell,f_j})x_{f_j} = [\beta]_{d_\ell} + \sum_{j=1}^{n-r}[\alpha_j]_{d_\ell}x_{f_j} = \left[\beta + \sum_{j=1}^{n-r} x_{f_j}\alpha_j \right]_{d_\ell}.$$

This tells us that the entries of the solution vector γ corresponding to lead variables (indices in D), are equal to those of a vector in the set S. We still need to check the other entries of the solution vector γ corresponding to the free variables (indices in F) to see if they are equal to the entries of the same vector in the set S. To this end, suppose $i = f_j \in F$. Then

$$\begin{aligned}
[\gamma]_i &= x_{f_j} = 0 + 0x_{f_1} + 0x_{f_2} + \cdots + 0x_{f_{j-1}} + 1x_{f_j} + 0x_{f_{j+1}} + \cdots + 0x_{f_{n-r}} \\
&= [\beta]_i + x_{f_1}[\alpha_1]_i + x_{f_2}[\alpha_2]_i + \cdots + x_{f_{j-1}}[\alpha_{j-1}]_i + x_{f_j}[\alpha_j]_i + x_{f_{j+1}}[\alpha_{j+1}]_i + \cdots + x_{f_{n-r}}[\alpha_{n-r}]_i \\
&= \left[\beta + \sum_{k=1}^{n-r} x_{f_k}\alpha_k\right]_i .
\end{aligned}$$

So $\gamma = \beta + \sum\limits_{k=1}^{n-r} x_{f_k}\alpha_k \in S$. □

6.4 Spanning set

Definition 6.4.1: Given a set of vectors

$$S = \{\alpha_1, \dots, \alpha_k\},$$

their *span*, $\langle S\rangle$, is the set of all possible linear combinations of $\alpha_1, \dots, \alpha_k$. Symbolically,

$$\langle S\rangle = \left\{ \sum_{i=1}^{k} a_i\alpha_i \;\middle|\; a_i \in \mathbb{R}, 1 \le i \le k \right\}.$$

For convenience, we let $\langle\varnothing\rangle = \{\mathbf{0}\}$. Also, $\langle\{\alpha_1, \dots, \alpha_k\}\rangle$ is simply denoted by $\langle\alpha_1, \dots, \alpha_k\rangle$.

Theorem 6.4.2: *Let $S = \{\alpha_1, \dots, \alpha_k\} \subset V = \mathbb{R}^m$. Then $\langle S\rangle$ is a subspace of V.*

Proof: Clearly, $\langle S\rangle$ is not empty, since $\mathbf{0}_m = \sum\limits_{i=1}^{k} 0\alpha_i \in \langle S\rangle$.

Main Questions:

1. Determine whether a vector $\alpha \in \langle S\rangle$.

2. Describe the set $\langle S\rangle$.

3. Is $\langle S\rangle = \mathbb{R}^m$ if S is a subset of $\mathbb{R}^m$?

Example 6.4.1: Consider the set $S \subset \mathbb{R}^4$

$$S = \{(1,1,3,1)^t, (2,1,2,-1)^t, (7,3,5,-5)^t, (1,1,-1,2)^t, (-1,0,9,0)^t\}.$$

Is $\alpha = (-15, -6, 19, 5)^t$ an element of $\langle S\rangle$?

We are asking whether there are scalars $a_1, a_2, a_3, a_4, a_5 \in \mathbb{R}$ such that

$$a_1\begin{pmatrix}1\\1\\3\\1\end{pmatrix} + a_2\begin{pmatrix}2\\1\\2\\-1\end{pmatrix} + a_3\begin{pmatrix}7\\3\\5\\-5\end{pmatrix} + a_4\begin{pmatrix}1\\1\\-1\\2\end{pmatrix} + a_5\begin{pmatrix}-1\\0\\9\\0\end{pmatrix} = \alpha = \begin{pmatrix}-15\\-6\\19\\5\end{pmatrix}.$$

Searching for these scalars is equivalent to finding solution to a linear system of equations with augmented matrix

$$\left(\begin{array}{ccccc|c} 1 & 2 & 7 & 1 & -1 & -15\\ 1 & 1 & 3 & 1 & 0 & -6\\ 3 & 2 & 5 & -1 & 9 & 19\\ 1 & -1 & -5 & 2 & 0 & 5 \end{array}\right)$$

which row-reduces to

$$\left(\begin{array}{ccccc|c} \textcircled{1} & 0 & -1 & 0 & 3 & 10\\ 0 & \textcircled{1} & 4 & 0 & -1 & -9\\ 0 & 0 & 0 & \textcircled{1} & -2 & -7\\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

At this point, we see that the system is consistent. So we know there is a solution for the five scalars a_1, a_2, a_3, a_4, a_5. This is enough evidence for us to say that $\alpha \in \langle S\rangle$. If we wished further evidence, we could compute an actual solution, say

$$a_1 = 2, a_2 = 1, a_3 = -2, a_4 = -3, a_5 = 2.$$

■

Example 6.4.2: Keeping the set S as in the previous example, is $\beta = (3, 1, 2, -1)^t$ an element of $\langle S\rangle$? We are asking whether there are scalars a_1, a_2, a_3, a_4, a_5 such that

$$a_1\begin{pmatrix}1\\1\\3\\1\end{pmatrix} + a_2\begin{pmatrix}2\\1\\2\\-1\end{pmatrix} + a_3\begin{pmatrix}7\\3\\5\\-5\end{pmatrix} + a_4\begin{pmatrix}1\\1\\-1\\2\end{pmatrix} + a_5\begin{pmatrix}-1\\0\\9\\0\end{pmatrix} = \beta = \begin{pmatrix}3\\1\\2\\-1\end{pmatrix}. \qquad (*)$$

This is equivalent to finding a solution to a linear system of equations with augmented matrix

$$\left(\begin{array}{ccccc|c} 1 & 2 & 7 & 1 & -1 & 3\\ 1 & 1 & 3 & 1 & 0 & 1\\ 3 & 2 & 5 & -1 & 9 & 2\\ 1 & -1 & -5 & 2 & 0 & -1 \end{array}\right)$$

which row-reduces to

$$\left(\begin{array}{ccccc|c} \textcircled{1} & 0 & -1 & 0 & 3 & 0\\ 0 & \textcircled{1} & 4 & 0 & -1 & 0\\ 0 & 0 & 0 & \textcircled{1} & -2 & 0\\ 0 & 0 & 0 & 0 & 0 & \textcircled{1} \end{array}\right)$$

At this point, we see that the system is inconsistent. So there are no scalars a_1, a_2, a_3, a_4, a_5 satisfying (*). Thus $\beta \notin \langle S\rangle$. ■

From the above illustration, we can easy to get the following theorem (proof is omitted):

Theorem 6.4.3: *Suppose $S = \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{R}^m$. Let A be an $m \times n$ matrix whose i-th column is α_i. A vector $\alpha \in \langle S \rangle$ if and only if $A\boldsymbol{x} = \alpha$ is consistent.*

Example 6.4.3: Let $\alpha = (v_1, v_2, v_3, v_4)^t \in \mathbb{R}^4$. Following Example 6.4.1, find the condition(s) of α such that $\alpha \in \langle S \rangle$.

Theorem 6.4.4: *Let $S = \{\alpha_1, \ldots, \alpha_m\} \subset \mathbb{R}^m$ and let A be the $m \times m$ matrix with the i-th column is α_i. A is non-singular if and only if $\langle S \rangle = \mathbb{R}^m$.*

Proof: Note that $\langle S \rangle \subseteq \mathbb{R}^m$ is always hold.

Theorem 6.4.5: *Suppose that $A \in M_{m,n}$ and $H = \text{rref}(A)$. Suppose that H has r leading columns, with indices given by $D = \{k_1, \ldots, k_r\}$, while the $n-r$ non-leading columns have indices $F = \{f_1, \ldots, f_{n-r}\}$. Construct the $n-r$ vectors α_j, $1 \le j \le n-r$, of length n,*

$$[\alpha_j]_i = \begin{cases} 1 & \text{if } i \in F,\ i = f_j \\ 0 & \text{if } i \in F,\ i \neq f_j \\ -[H]_{k,f_j} & \text{if } i \in D,\ i = d_k \end{cases}$$

Then the null space of A is given by

$$\mathcal{N}(A) = \langle \alpha_1, \ldots, \alpha_{n-r} \rangle .$$

Proof: This theorem follows from Theorem 6.3.3 by letting $\boldsymbol{b} = \mathbf{0}_m$. $\square$

Do not memorize this theorem. Instead, study the examples below.

Example 6.4.4: Find a spanning set of $\mathcal{N}(A)$, where

$$A = \begin{pmatrix} 2 & 1 & 5 & 1 & 5 & 1 \\ 1 & 1 & 3 & 1 & 6 & -1 \\ -1 & 1 & -1 & 0 & 4 & -3 \\ -3 & 2 & -4 & -4 & -7 & 0 \\ 3 & -1 & 5 & 2 & 2 & 3 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} \textcircled{1} & 0 & 2 & 0 & -1 & 2 \\ 0 & \textcircled{1} & 1 & 0 & 3 & -1 \\ 0 & 0 & 0 & \textcircled{1} & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$