Chapter 7: Linear Independence, Column and Row Spaces

7.1 Span by Fewer Vectors

Example 7.1.1: Let $\alpha_1 = (2,-3,1)^t$, $\alpha_2 = (1,4,1)^t$, $\alpha_3 = (7,-5,4)^t$ and $\alpha_4 = (-7,-6,-5)^t$. Let $W = \langle \alpha_1, \alpha_2, \alpha_3, \alpha_4 \rangle$.

Let

$$D = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 7 & -7 \\ -3 & 4 & -5 & -6 \\ 1 & 1 & 4 & -5 \end{pmatrix}.$$

Check that the vector $\beta = (2,3,0,1)^t$ is a solution to the homogeneous system $D\boldsymbol{x} = \mathbf{0}_3$. That is,

$$2\alpha_1 + 3\alpha_2 + 0\alpha_3 + 1\alpha_4 = \mathbf{0}_3.$$

We may rewrite it as

$$\alpha_4 = (-2)\alpha_1 + (-3)\alpha_2.$$

This equation says that whenever we encounter the vector α_4, we can replace it with a specific linear combination of the vectors α_1 and α_2. So using α_4 in the spanning set of W along with α_1 and α_2 is excessive.

Since any linear combination of $\alpha_1, \alpha_2, \alpha_3, \alpha_4$

$$\begin{aligned} & a\alpha_1 + b\alpha_2 + c\alpha_3 + d\alpha_4 \\ = {} & a\alpha_1 + b\alpha_2 + c\alpha_3 + d((-2)\alpha_1 + (-3)\alpha_2) \\ = {} & (a-2d)\alpha_1 + (b-3d)\alpha_2 + c\alpha_3 \end{aligned}$$

can be rewritten as a linear combination of $\alpha_1, \alpha_2, \alpha_3$. So $W = \langle \alpha_1, \alpha_2, \alpha_3 \rangle$.

So the span of our set of vectors, W, has not changed, but we have *described* it by the span of a set of *three* vectors, rather than *four*. Furthermore, we can achieve yet another, similar, reduction.

Check that the vector

$$\gamma = (-3,-1,1,0)^t$$

is a solution to the homogeneous system $D\boldsymbol{x} = \mathbf{0}_3$. We can write the linear combination,

$$(-3)\alpha_1 + (-1)\alpha_2 + 1\alpha_3 = \mathbf{0}_3.$$

We can solve for α_3,

$$\alpha_3 = 3\alpha_1 + 1\alpha_2.$$

This equation says that whenever we encounter the vector α_3, we can replace it with a specific linear combination of the vectors α_1 and α_2. So, as before, the vector α_3 is not needed in the description of W, provided we have α_1 and α_2 available. So

$$W = \langle \alpha_1, \alpha_2 \rangle.$$

From the above equation, we may also obtain $\alpha_2 = -3\alpha_1 + \alpha_3$ and $\alpha_1 = -\frac{1}{3}\alpha_2 + \frac{1}{3}\alpha_3$. So we may get that $W = \langle \alpha_1, \alpha_3 \rangle = \langle \alpha_2, \alpha_3 \rangle$. ■

So W began life as the span of a set of four vectors, and we have now shown (utilizing solutions to a homogeneous system) that W can also be described as the span of a set of just two vectors. Convince yourself that we cannot go any further. In other words, it is not possible to dismiss either α_1 or α_2 in a similar fashion and winnow the set down to just one vector. What was it about the original set of four vectors that allowed us to declare certain vectors as surplus? And just which vectors were we able to dismiss? And why did we have to stop once we had two vectors remaining? The answers to these questions motivate **linear independence** our next section and next definition, and so are worth considering carefully now.

7.2 Linearly Independent

Definition 7.2.1: Given a set of vectors $S = \{\alpha_1, \dots, \alpha_n\}$, a true statement of the form

$$a_1\alpha_1 + \cdots + a_n\alpha_n = \mathbf{0}$$

is a *relation of linear dependence* (or *linear relation*) on S. If this statement is formed in a trivial fashion, i.e., $a_i = 0$, $1 \le i \le n$, then we say it is the *trivial relation of linear dependence* on S.

Definition 7.2.2: The set of vectors $S = \{\alpha_1, \dots, \alpha_n\}$ is *linearly dependent* if there is a relation of linear dependence on S that is not trivial. We also say that the vectors $\alpha_1, \dots, \alpha_n$ are *linearly dependent*. In the case where the *only* relation of linear dependence on S is the trivial one, then S is a *linearly independent* set of vectors. We also say that the vectors $\alpha_1, \dots, \alpha_n$ are *linearly independent*.

Remark 7.2.3: In other word, $\alpha_1, \dots, \alpha_n$ are linearly dependent if (and only if) there are $a_1, \dots, a_n \in \mathbb{R}$, not all zero, such that $\sum\limits_{i=1}^{n} a_i\alpha_i = \mathbf{0}$.

Remark 7.2.4: To prove $\alpha_1, \dots, \alpha_n$ are linearly independent, we need to start with a relation of linear dependence and somehow conclude that the scalars involved *must all be zero*, i.e., that the relation of linear dependence only happens in the trivial fashion. In mathematical (symbolic) statement

$$\boxed{a_1\alpha_1 + \cdots + a_n\alpha_n = \sum_{i=1}^{n} a_i\alpha_i = \mathbf{0} \Longrightarrow a_i = 0\ \forall i.}$$

Example 7.2.1: Consider the set of $n = 4$ vectors from $\mathbb{R}^5$,

$$S = \{(2, -1, 3, 1, 2)^t, (1, 2, -1, 5, 2)^t, (2, 1, -3, 6, 1)^t, (-6, 7, -1, 0, 1)^t\}.$$

To determine linear independence we first form a relation of linear dependence,

$$a_1 \begin{pmatrix} 2 \\ -1 \\ 3 \\ 1 \\ 2 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 5 \\ 2 \end{pmatrix} + a_3 \begin{pmatrix} 2 \\ 1 \\ -3 \\ 6 \\ 1 \end{pmatrix} + a_4 \begin{pmatrix} -6 \\ 7 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \mathbf{0}.$$

We know that $a_1 = a_2 = a_3 = a_4 = 0$ is a solution to this equation, but that is of no interest whatsoever. That is *always* the case, no matter what four vectors we might have chosen. We are curious to know if there are other, nontrivial, solutions. Row-reducing

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

We could solve this homogeneous system completely, but for this example all we need is one nontrivial solution.

Example 7.2.2: Consider the set of $n = 4$ vectors from $\mathbb{R}^5$,

$$T = \{(2, -1, 3, 1, 2)^t, (1, 2, -1, 5, 2)^t, (2, 1, -3, 6, 1)^t, (-6, 7, -1, 1, 1)^t\}.$$

To determine linear independence we first form a relation of linear dependence,

$$a_1 \begin{pmatrix} 2 \\ -1 \\ 3 \\ 1 \\ 2 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 5 \\ 2 \end{pmatrix} + a_3 \begin{pmatrix} 2 \\ 1 \\ -3 \\ 6 \\ 1 \end{pmatrix} + a_4 \begin{pmatrix} -6 \\ 7 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{0}.$$

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

From the form of this matrix, we see that there are no free variables. So the solution is unique, and because the system is homogeneous, this unique solution is the trivial solution. So we know that there is only one way to combine the four vectors of T into a relation of linear dependence. And that one way is the easy and obvious way. In this situation we say that T is linearly independent. ■

Theorem 7.2.5: *Suppose that $S = \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{R}^m$ and A is the $m \times n$ matrix whose columns are the vectors in S. The following statements are equivalent:*

(1) S is a linearly independent set,

(2) the homogeneous system $A\boldsymbol{x} = \mathbf{0}_m$ has a unique solution,

(3) $r(A) = \operatorname{rank}(A) = n$.

Proof:

Example 7.2.3: Is the set of vectors

$$S = \{(2, -1, 3, 1, 0, 3)^t, (9, -6, -2, 3, 2, 1)^t, (1, 1, 1, 0, 0, 1)^t, (-3, 1, 4, 2, 1, 2)^t, (6, -2, 1, 4, 3, 2)^t\}$$

linearly independent or linearly dependent?

Solution: Theorem 7.2.5 suggests we place these vectors into a matrix as columns and analyze the row-reduced version of the matrix,

$$\begin{pmatrix} 2 & 9 & 1 & -3 & 6 \\ -1 & -6 & 1 & 1 & -2 \\ 3 & -2 & 1 & 4 & 1 \\ 1 & 3 & 0 & 2 & 4 \\ 0 & 2 & 0 & 1 & 3 \\ 3 & 1 & 1 & 2 & 2 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} \textcircled{1} & 0 & 0 & 0 & -1 \\ 0 & \textcircled{1} & 0 & 0 & 1 \\ 0 & 0 & \textcircled{1} & 0 & 2 \\ 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now we have $r = 4 < 5 = n$. By Theorem 7.2.5, S is a linearly dependent set. ■

Example 7.2.4: Consider $n = 9$ vectors from $\mathbb{R}^4$
$(-1, 3, 1, 2)^t$, $(7, 1, -3, 6)^t$, $(1, 2, -1, -2)^t$, $(0, 4, 2, 9)^t$, $(5, -2, 4, 3)^t$, $(2, 1, -6, 4)^t$, $(3, 0, -3, 1)^t$, $(1, 1, 5, 3)^t$, $(-6, -1, 1, 1)^t$.

To employ Theorem 7.2.5, we form a 4×9 matrix, C, whose columns are these vectors

$$C = \begin{pmatrix} -1 & 7 & 1 & 0 & 5 & 2 & 3 & 1 & -6 \\ 3 & 1 & 2 & 4 & -2 & 1 & 0 & 1 & -1 \\ 1 & -3 & -1 & 2 & 4 & -6 & -3 & 5 & 1 \\ 2 & 6 & -2 & 9 & 3 & 4 & 1 & 3 & 1 \end{pmatrix}.$$

Theorem 7.2.6: *Suppose that $S = \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{R}^m$ and $n > m$. Then S is a linearly dependent set.*

Proof:

We will now specialize to sets of n vectors from $\mathbb{R}^n$. This will put Theorem 7.2.6 off-limits, while Theorem 7.2.5 will involve square matrices.

Theorem 7.2.7: *Suppose that A is a square matrix. Then A is nonsingular if and only if the columns of A form a linearly independent set.*

Proof:

$$\begin{aligned} A \text{ is nonsingular} &\iff A\boldsymbol{x} = \mathbf{0} \text{ has a unique solution} \\ &\iff \text{columns of } A \text{ are linearly independent.} \end{aligned}$$

□

Combining Theorem 6.4.4 and the above theorem, here is the update to Theorem 5.3.6.

Theorem 7.2.8: *Suppose that $A \in M_n$. The following are equivalent.*

1. *A is nonsingular.*

2. *A is row equivalent to I_n.*

3. *$\mathcal{N}(A) = \{\mathbf{0}_n\}$.*

4. *The linear system $\mathcal{LS}(A, \boldsymbol{b})$ has a unique solution for every possible choice of $\boldsymbol{b}$.*

5. *A is invertible.* (Skip it if Chapter 5 has not been taught.)

6. *The columns of A form a linearly independent set.*

7. *The columns of A span $\mathbb{R}^n$.*

We update to Theorem 6.4.5

Theorem 7.2.9: *Suppose that $A \in M_{m,n}$ and $H = \mathrm{rref}(A)$. Suppose that H has r leading columns, with indices given by $D = \{d_1, \ldots, d_r\}$, 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,*

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

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

(2) $\alpha_1, \ldots, \alpha_{n-r}$ *are linearly independent.*

Proof: (1) was proved in Theorem 6.4.5.

To prove (2), we start with $\sum_{i=1}^{n-r} a_i\alpha_i = \mathbf{0}$ for some $a_i \in \mathbb{R}$ (see Remark 7.2.4). Note that $[\alpha_j]_i = \delta_{f_j,i}$ for $i \in F$.

Example 7.2.5: Find the null space of

$$A = \begin{pmatrix} -2 & -1 & -2 & -4 & 4 \\ -6 & -5 & -4 & -4 & 6 \\ 10 & 7 & 7 & 10 & -13 \\ -7 & -5 & -6 & -9 & 10 \\ -4 & -3 & -4 & -6 & 6 \end{pmatrix}$$

Solution:

$$A \xrightarrow{rref} \begin{pmatrix} \textcircled{1} & 0 & 0 & 1 & -2 \\ 0 & \textcircled{1} & 0 & -2 & 2 \\ 0 & 0 & \textcircled{1} & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

x_4 and x_5 are free variables.
α_1 corresponding to $x_4 = 1$, $x_5 = 0$ and α_2 corresponding to $x_4 = 0$, $x_5 = 1$. We have

$$\alpha_1 = \begin{pmatrix} -1 \\ 2 \\ -2 \\ 1 \\ 0 \end{pmatrix} \quad \alpha_2 = \begin{pmatrix} 2 \\ -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

By Theorem 7.2.9, $\mathcal{N}(A) = \langle \alpha_1, \alpha_2 \rangle$. ■

Suppose a set contains a zero vector, say $S = \{\mathbf{0}, \alpha_2, \ldots, \alpha_n\}$. Then

$$1\mathbf{0} + 0\alpha_2 + \cdots + 0\alpha_n = \mathbf{0}.$$

Hence S is linearly dependent.

So, we only consider some finite sets containing non-zero vectors.

Theorem 7.2.10: *A set of non-zero vectors* $\{\alpha_1, \ldots, \alpha_n\}$ *is linearly dependent if and only if there is a vector* α_k *that is a linear combination of* $\alpha_1, \alpha_2, \ldots, \alpha_j$ *with* $j < k$.

Proof:

7.3 Casting-out Method and Column Space

In Example 7.1.1 we used four vectors to create a span. With a relation of linear dependence in hand, we were able to *cast out* two of these four vectors and create the same span from a subset of just two vectors from the original set of four. We did have to take some care as to just which vector we casted out. In the next example, we will be more methodical about just how we choose to eliminate vectors from a linearly dependent set while preserving a span. This method is called the *casting-out method*.

Example 7.3.1: We begin with a set S containing seven vectors from $\mathbb{R}^4$,

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 0 \\ -4 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ -3 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 9 \\ -4 \\ 8 \end{pmatrix}, \begin{pmatrix} 7 \\ -13 \\ 12 \\ -31 \end{pmatrix}, \begin{pmatrix} -9 \\ 7 \\ -8 \\ 37 \end{pmatrix} \right\}$$

and define $W = \langle S \rangle$.

The set S is obviously linearly dependent, since we have $n = 7$ vectors from $\mathbb{R}^4$. So we can slim down S some, and still create W as the span of a smaller set of vectors.

As a device for identifying linear relations among the vectors of S, we place the seven column vectors of S into a matrix as columns,

$$A = \begin{pmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{pmatrix}$$

A nontrivial solution to $A\boldsymbol{x} = \mathbf{0}$ will give us a nontrivial linear relation on the columns of A (which are the elements of the set S). The rref of A is the matrix

$$H = \begin{pmatrix} \textcircled{1} & 4 & 0 & 0 & 2 & 1 & -3 \\ 0 & 0 & \textcircled{1} & 0 & 1 & -3 & 5 \\ 0 & 0 & 0 & \textcircled{1} & 2 & -6 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

So we can easily create solutions to the homogeneous system $A\boldsymbol{x} = \mathbf{0}$ using the free variables x_2, x_5, x_6, x_7. Any such solution will provide a relation of linear dependence on the columns of A. These solutions will allow us to solve for one column vector as a linear combination of some others, in the spirit of Theorem 7.2.10, and remove that vector from the set. We will set about forming these linear combinations methodically.

Set the free variable $x_2 = 1$, and set the other free variables to zero. Then a solution to $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ (also to $\boldsymbol{H}\boldsymbol{x} = \mathbf{0}$) is

$$\alpha_1 = (-4, 1, 0, 0, 0, 0, 0)^t$$

which can be used to create the linear combination

$$(-4)A_{*1} + 1A_{*2} + 0A_{*3} + 0A_{*4} + 0A_{*5} + 0A_{*6} + 0A_{*7} = \mathbf{0}.$$

Then A_{*2} can be expressed as a linear combination of $\{A_{*1}\}$,

$$A_{*2} = 4A_{*1}.$$

This means that A_{*2} is surplus, and we can span W just as well with a smaller set with this vector removed,

$$W = \langle A_{*1}, A_{*3}, A_{*4}, A_{*5}, A_{*6}, A_{*7} \rangle .$$

Now, set the free variable $x_5 = 1$, and set the other free variables to zero. Then a solution to $\boldsymbol{H}\boldsymbol{x} = \mathbf{0}$ is

$$\alpha_2 = (-2, 0, -1, -2, 1, 0, 0)^t$$

which can be used to create the linear combination

$$(-2)A_{*1} + 0A_{*2} + (-1)A_{*3} + (-2)A_{*4} + 1A_{*5} + 0A_{*6} + 0A_{*7} = \mathbf{0}.$$

Then A_{*5} can be expressed as a linear combination of $\{A_{*1}, A_{*3}, A_{*4}\}$,

$$A_{*5} = 2A_{*1} + 1A_{*3} + 2A_{*4}.$$

This means that A_{*5} is surplus, and we can span W just as well with a smaller set with this vector removed,

$$W = \langle A_{*1}, A_{*3}, A_{*4}, A_{*6}, A_{*7} \rangle .$$

Do it again, set the free variable $x_6 = 1$, and set the other free variables to zero. Then we have

$$\alpha_3 = (-1, 0, 3, 6, 0, 1, 0)^t$$

which can be used to create the linear combination

$$(-1)A_{*1} + 0A_{*2} + 3A_{*3} + 6A_{*4} + 0A_{*5} + 1A_{*6} + 0A_{*7} = \mathbf{0}.$$

Then A_{*6} can be expressed as a linear combination of $\{A_{*1}, A_{*3}, A_{*4}\}$,

$$A_{*6} = 1A_{*1} + (-3)A_{*3} + (-6)A_{*4}.$$

This means that A_{*6} is surplus, and we can span W just as well with a smaller set with this vector removed,

$$W = \langle A_{*1}, A_{*3}, A_{*4}, A_{*7} \rangle .$$

Set the free variable $x_7 = 1$, and set the other free variables to zero. We have

$$\alpha_4 = (3, 0, -5, -6, 0, 0, 1)^t$$

which can be used to create the linear combination

$$3A_{*1} + 0A_{*2} + (-5)A_{*3} + (-6)A_{*4} + 0A_{*5} + 0A_{*6} + 1A_{*7} = \mathbf{0}$$

Then A_{*7} can be expressed as a linear combination of $\{A_{*1}, A_{*3}, A_{*4}\}$,

$$A_{*7} = (-3)A_{*1} + 5A_{*3} + 6A_{*4}.$$

This means that A_{*7} is surplus, and we can span W just as well with a smaller set with this vector removed,

$$W = \langle A_{*1}, A_{*3}, A_{*4} \rangle .$$

■

You might think we could keep this up, but we have run out of free variables. And not coincidentally, the set $\{A_{*1}, A_{*3}, A_{*4}\}$ is linearly independent (check this!). It should be clear how each free variable was used to eliminate a column from the set used to span the column space, as this will be the essence of the proof of the next theorem.

Definition 7.3.1: Let $A \in M_{m,n}$. A column of A corresponding the leading column of rref(A) is called a *leading column of A*. The leading column index of rref(A) is also the *leading column index of A*.

Important: The above example shows that

1. The leading columns of A form a linearly independent set.
   The leading column index of A is $D = \{1, 3, 4\}$. So $\{A_{*1}, A_{*3}, A_{*4}\}$ is a linearly independent sets.

2. All the other columns of A are linear combinations of A_{*1}, A_{*3}, A_{*4}.
   In fact, the relation can be written explicitly. First, obviously by H

$$\begin{aligned}
H_{*2} &= 4H_{*1} = 4\boldsymbol{e}_1 \\
H_{*5} &= 2H_{*1} + H_{*3} + 2H_{*4} = 2\boldsymbol{e}_1 + \boldsymbol{e}_2 + 2\boldsymbol{e}_3 \\
H_{*6} &= H_{*1} - 3H_{*3} - 6H_{*4} = \boldsymbol{e}_1 - 3\boldsymbol{e}_2 - 6\boldsymbol{e}_3 \\
H_{*7} &= -3H_{*1} + 5H_{*3} + 6H_{*4} = -3\boldsymbol{e}_1 + 5\boldsymbol{e}_2 + 6\boldsymbol{e}_3
\end{aligned}$$

   Correspondingly we have

$$\begin{aligned}
A_{*2} &= 4A_{*1} \\
A_{*5} &= 2A_{*1} + A_{*3} + 2A_{*4} \\
A_{*6} &= A_{*1} - 3A_{*3} - 6A_{*4} \\
A_{*7} &= -3A_{*1} + 5A_{*3} + 6A_{*4}
\end{aligned}$$

Suppose $A \in M_{m,n}$. Let $\sum_{i=1}^{n} a_i A_{*i}$ be a linear combination of A_{*i}'s, where $a_i \in \mathbb{R}$. Let P be any invertible matrix. Then

$$\mathbf{0}_m = \sum_{i=1}^{n} a_i A_{*i} \iff \mathbf{0}_m = P\left(\sum_{i=1}^{n} a_i A_{*i}\right) = \sum_{i=1}^{n} a_i (PA_{*i}). \tag{7.1}$$

In particular, if P is such that $PA = \text{rref}(A)$, then PA_{*i} is the i-th column of rref(A). So their linear relation can be easily inspected.

Definition 7.3.2: Suppose $A \in M_{m,n}$. The subspace spanned by the columns of A is called the *column space of* A and denoted by $\mathcal{C}(A)$. That is, $\mathcal{C}(A) = \langle A_{*1}, \ldots, A_{*n} \rangle$.

Now, we restate Theorem 6.3.2 as

Theorem 7.3.3: *Suppose $A \in M_{m,n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then $\boldsymbol{b} \in \mathcal{C}(A)$ if and only if $\mathcal{LS}(A, \boldsymbol{b})$ is consistent.*

Thus, an alternative (and popular) definition of the column space of an $m \times n$ matrix A is

$$\mathcal{C}(A) = \{\boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{y} = A\boldsymbol{x} \text{ for some } \boldsymbol{x} \in \mathbb{R}^n\} = \{A\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$

Lemma 7.3.4: *Let $\{\alpha_1, \ldots, \alpha_s\}$ and $\{\beta_1, \ldots, \beta_t\}$ be two subsets of a vector space. If each β_j is a linear combination of $\{\alpha_1, \ldots, \alpha_s\}$, then $\langle \beta_1, \ldots, \beta_t \rangle \subseteq \langle \alpha_1, \ldots, \alpha_s \rangle$.*

Proof:

Theorem 7.3.5: *Suppose $A = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix} \in M_{m,n}$. Let $H = \operatorname{rref}(A)$ with $D = \{d_1, \ldots, d_r\}$ the set of indices for the leading columns of H (also of A). Then*

1. *$T = \{\alpha_{d_1}, \ldots, \alpha_{d_r}\}$ is a linearly independent set.*
2. *$\mathcal{C}(A) = \langle T \rangle$.*

Proof: Note that $H_{*d_j} = \boldsymbol{e}_j \in \mathbb{R}^m$, $1 \le j \le r$.

Example 7.3.2: Let $S = \{\alpha_1 = (1,0,-1,1)^t, \alpha_2 = (0,1,2,-1)^t, \alpha_3 = (1,1,1,0)^t, \alpha_4 = (-1,1,1,2)^t, \alpha_5 = (-2,3,2,7)^t\}$. Find a maximal linearly independent subset of S (i.e., the largest and linearly independent subset of S).

Since the rref of a matrix is unique, the procedure of Theorem 7.3.5 leads us to a unique set T. However, there is a wide variety of possibilities for sets T that are linearly independent and which can be employed in a span to span $\mathcal{C}(A)$. Without proof, we list two other possibilities for the above example:

$$T' = \{\alpha_1, \alpha_3, \alpha_4\} \text{ and } T^* = \{\alpha_2, \alpha_3, \alpha_5\}.$$

Can you prove that T' and T^* are linearly independent sets and $\mathcal{C}(A) = \langle T' \rangle = \langle T^* \rangle$?

These are maximal linear independent subsets of S too.

Example 7.3.3: Let

$$A = \begin{pmatrix} 1 & 2 & 7 & 1 & -1 \\ 1 & 1 & 3 & 1 & 0 \\ 3 & 2 & 5 & -1 & 9 \\ 1 & -1 & -5 & 2 & 0 \end{pmatrix}.$$

Find $\mathcal{C}(A)$ as a null space of a linear system or a null space of some matrix.

Solution:

Example 7.3.4: Let

$$A = \begin{pmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{pmatrix}.$$

Find a minimal subset of the set of columns of A that spans $\mathcal{C}(A)$ (this is called a minimal spanning set of $\mathcal{C}(A)$, or a basis of $\mathcal{C}(A)$ later).

Solution: This is the same matrix in Example 7.3.1.

Restate the last statement of Theorem 7.2.8 we have

Theorem 7.3.6: *Suppose $A \in M_n$. A is nonsingular if and only if $\mathcal{C}(A) = \mathbb{R}^n$.*

7.4 Row Space

Definition 7.4.1: Suppose $A \in M_{m,n}$. The *row space* of A, $\mathcal{R}(A)$, is the column space of A^t, i.e., $\mathcal{R}(A) = \mathcal{C}(A^t)$.

Informally, the row space is the set of all linear combinations of the rows of A. However, we write the rows as column vectors, thus the necessity of using the transpose to make the rows into columns. Additionally, with the row space defined in terms of the column space, all of the previous results of this chapter can be applied to row spaces.

Example 7.4.1: Find $\mathcal{R}(A)$ for

$$A = \begin{pmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{pmatrix}.$$

To build the row space, we transpose the matrix,

$$A^t = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 4 & 8 & 0 & -4 \\ 0 & -1 & 2 & 2 \\ -1 & 3 & -3 & 4 \\ 0 & 9 & -4 & 8 \\ 7 & -13 & 12 & -31 \\ -9 & 7 & -8 & 37 \end{pmatrix}$$

Then the columns of this matrix are used in a span to build the row space,

$$\mathcal{R}(A) = \mathcal{C}(A^t) = \left\langle \begin{pmatrix} 1 \\ 4 \\ 0 \\ -1 \\ 0 \\ 7 \\ -9 \end{pmatrix}, \begin{pmatrix} 2 \\ 8 \\ -1 \\ 3 \\ 9 \\ -13 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ -3 \\ -4 \\ 12 \\ -8 \end{pmatrix}, \begin{pmatrix} -1 \\ -4 \\ 2 \\ 4 \\ 8 \\ -31 \\ 37 \end{pmatrix} \right\rangle.$$

First, row-reduce A^t,

$$\begin{pmatrix} \textcircled{1} & 0 & 0 & -\frac{31}{7} \\ 0 & \textcircled{1} & 0 & \frac{12}{7} \\ 0 & 0 & \textcircled{1} & \frac{13}{7} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since the leading columns have indices $D = \{1, 2, 3\}$, the column space of A^t can be spanned by the first three columns of A^t,

$$\mathcal{R}(A) = \mathcal{C}(A^t) = \left\langle \begin{pmatrix} 1 \\ 4 \\ 0 \\ -1 \\ 0 \\ 7 \\ -9 \end{pmatrix}, \begin{pmatrix} 2 \\ 8 \\ -1 \\ 3 \\ 9 \\ -13 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ -3 \\ -4 \\ 12 \\ -8 \end{pmatrix} \right\rangle.$$

■

Theorem 7.4.2: *Suppose A and B are row-equivalent $m \times n$ matrices. Then $\mathcal{R}(A) = \mathcal{R}(B)$.*

Proof:

Theorem 7.4.3: *Suppose that A is a matrix and $H = \mathrm{rref}(A)$. Let S be the set of nonzero columns of H^t. Then*

1. *$\mathcal{R}(A) = \langle S \rangle$.*
2. *S is a linearly independent set.*

Proof:

Example 7.4.2: Let $X = \left\langle (1,2,1,6,6)^t,\ (3,-1,2,-1,6)^t,\ (1,-1,0,-1,-2)^t,\ (-3,2,-3,6,-10)^t \right\rangle$.

Let A be the matrix whose rows are the vectors in X, so by design $X = \mathcal{R}(A)$. Now

$$A = \begin{pmatrix} 1 & 2 & 1 & 6 & 6 \\ 3 & -1 & 2 & -1 & 6 \\ 1 & -1 & 0 & -1 & -2 \\ -3 & 2 & -3 & 6 & -10 \end{pmatrix}.$$

We get

$$H = \text{rref}(A) = \begin{pmatrix} \textcircled{1} & 0 & 0 & 2 & -1 \\ 0 & \textcircled{1} & 0 & 3 & 1 \\ 0 & 0 & \textcircled{1} & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then the above theorem says that we can grab the nonzero columns of H^t and write

$$X = \mathcal{R}(A) = \mathcal{R}(H) = \left\langle (1,0,0,2,-1)^t, (0,1,0,3,1)^t, (0,0,1,-2,5)^t \right\rangle.$$

Note that, the vectors in the spanning set of X here are not come form the original vectors. ■

Theorem 7.4.4: *Suppose A is a matrix. Then $\mathcal{C}(A) = \mathcal{R}(A^t)$.*

Proof: $\mathcal{C}(A) = \mathcal{C}((A^t)^t) = \mathcal{R}(A^t)$. □

Example 7.4.3: Find a spanning set of the column space of A in Example 7.4.1. Here

$$A = \begin{pmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{pmatrix}.$$

Method 1.

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

By Theorem 7.3.5

$$\mathcal{C}(A) = \left\langle (1,2,0,-1)^t, (0,-1,2,2)^t, (-1,3,-3,4)^t \right\rangle.$$

Method 2. The transpose of A is

$$A^t = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 4 & 8 & 0 & -4 \\ 0 & -1 & 2 & 2 \\ -1 & 3 & -3 & 4 \\ 0 & 9 & -4 & 8 \\ 7 & -13 & 12 & -31 \\ -9 & 7 & -8 & 37 \end{pmatrix}.$$

We have

$$H = \text{rref}(A^t) = \begin{pmatrix} \textcircled{1} & 0 & 0 & -\frac{31}{7} \\ 0 & \textcircled{1} & 0 & \frac{12}{7} \\ 0 & 0 & \textcircled{1} & \frac{13}{7} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence

$$\mathcal{C}(A) \xlongequal{\text{Theorem 7.4.4}} \mathcal{R}(A^t) \xlongequal{\text{Theorem 7.4.2}} \mathcal{R}(H) \xlongequal{\text{Definition}} \mathcal{C}(H^t) = \left\langle \begin{bmatrix} 1 \\ 0 \\ 0 \\ -\frac{31}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{12}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{13}{7} \end{bmatrix} \right\rangle.$$

■

This is a very nice description of the column space. Fewer vectors than the 7 involved in the definition, and the pattern of the zeros and ones in the first 3 slots can be used to advantage. For example, let us check if

$$\beta = (3, 9, 1, 4)^t$$

is in $\mathcal{C}(A)$ or not. If it is, then

$$\beta = \begin{bmatrix} 3 \\ 9 \\ 1 \\ 4 \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ 0 \\ -\frac{31}{7} \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{12}{7} \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{13}{7} \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ -\frac{31}{7}x + \frac{12}{7}y + \frac{13}{7}z \end{bmatrix},$$

for some $x, y, z \in \mathbb{R}$.

From the first three coordinate, we have $x = 3, y = 9, z = 1$. Let us check the last coordinate:

$$-\frac{31}{7} \times 3 + \frac{12}{7} \times 9 + \frac{13}{7} \times 1 = 4.$$

Hence $\beta \in \mathcal{C}(A)$.

Remark 7.4.5: Both methods describe algorithms to find bases for the column space (i.e., linear independent sets generate the column space which will be introduced in next chapter). Here are the differences.

1. In Method 1, we find a subset of columns that form a basis. However in Method 2, the basis is not a subset of columns.

2. Given a vector $\beta \in \mathcal{C}(A)$, it is easier to express it as a linear combination of the basis given by Method 2.