Chapter 5

Vector Spaces: Theory and Practice

So far, we have worked with vectors of length n and performed basic operations on them like scaling and addition. Next, we looked at solving linear systems via Gaussian elimination and LU factorization. Already, we ran into the problem of what to do if a zero "pivot" is encountered. What if this cannot be fixed by swapping rows? Under what circumstances will a linear system not have a solution? When will it have more than one solution? How can we describe the set of all solutions? To answer these questions, we need to dive deeper into the theory of linear algebra.

5.1 Vector Spaces

The reader should be quite comfortable with the simplest of vector spaces: $\mathbb{R}$,$\mathbb{R}^2$, and R^3, which represent the points in one-dimentional, two-dimensional, and three-dimensional (real valued) space, respectively. A vector $x \in \mathbb{R}^n$ is represented by the column of n real numbers $x = \begin{pmatrix} \chi_0 \\ \vdots \\ \chi_{n-1} \end{pmatrix}$ which one can think of as the direction (vector) from the origin (the point $0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$) to the point $x = \begin{pmatrix} \chi_0 \\ \vdots \\ \chi_{n-1} \end{pmatrix}$. However, notice that a direction is position independent: You can think of it as a direction anchored anywhere in $\mathbb{R}^n$.

What makes the set of all vectors in $\mathbb{R}^n$ a space is the fact that there is a scaling operation (multiplication by a scalar) defined on it, as well as the addition of two vectors: The definition of a space is a set of elements (in this case vectors in $\mathbb{R}^n$) together with addition and multiplication (scaling) operations such that (1) for any two elements in the set, the element that results from adding these elements is also in the set; (2) scaling an element in the set results in an element in the set; and (3) there is an element in the set, denoted by 0, such that additing it to another element results in that element and scaling by 0 results in the 0 element.

Example 5.1 Let $x, y \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$. Then

- $z = x + y \in \mathbb{R}^2$;
- $\alpha \cdot x = \alpha x \in \mathbb{R}^2$; and
- $0 \in \mathbb{R}^2$ and $0 \cdot x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

In this document we will talk about *vector spaces* because the spaces have vectors as their elements.

Example 5.2 Consider the set of all real valued $m \times n$ matrices, $\mathbb{R}^{m \times n}$. Together with matrix addition and multiplication by a scalar, this set is a vector space.

Note that an easy way to visualize this is to take the matrix and view it as a vector of length $m \cdot n$.

Example 5.3 Not all spaces are vector spaces. For example, the spaces of all functions defined from $\mathbb{R}$ to $\mathbb{R}$ has addition and multiplication by a scalar defined on it, but it is not a vectors space. (It is a space of functions instead.)

Recall the concept of a subset, B, of a given set, A. All elements in B are elements in A. If A is a vector space we can ask ourselves the question of when B is also a vector space. The answer is that B is a vector space if (1) $x, y \in B$ implies that $x + y \in B$; (2) $x \in B$ and $\alpha \in B$ implies $\alpha x \in B$; and (3) $0 \in B$ (the zero vector). We call a subset of a vector space that is also a vector space a *subspace*.

Example 5.4 Reason that one does not need to explicitly say that the zero vector is in a (sub)space.

Definition 5.5 Let A be a vector space and let B be a subset of A. Then B is a subspace of A if (1) $x, y \in B$ implies that $x + y \in B$; and (2) $x \in B$ and $\alpha \in R$ implies that $\alpha x \in B$.

One way to describe a subspace is that it is a subset that is *closed* under addition and scalar multiplication.

Example 5.6 The empty set is a subset of $\mathbb{R}^n$. Is it a subspace? Why?

Exercise 5.7 What is the smallest subspace of $\mathbb{R}^n$? (Smallest in terms of the number of elements in the subspace.)

5.2 Why Should We Care?

Example 5.8 Consider

$$A = \begin{pmatrix} 3 & -1 & 2 \\ 1 & 2 & 0 \\ 4 & 1 & 2 \end{pmatrix}, \quad b_0 = \begin{pmatrix} 8 \\ -1 \\ 7 \end{pmatrix}, \quad \text{and} \quad b_1 = \begin{pmatrix} 5 \\ -1 \\ 7 \end{pmatrix}$$

Does $Ax = b_0$ have a solution? The answer is yes: $x = (1, -1, 2)^T$. Does $Ax = b_1$ have a solution? The answer is no. Does $Ax = b_0$ have any other solutions? The answer is yes.

The above example motivates the question of when a linear system has a solution, when it doesn't, and how many solutions it has. We will try to answer that question in this section.

Let $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $Ax = b$. Partition

$$A \rightarrow \left(\begin{array}{c|c|c|c} a_0 & a_1 & \cdots & a_{n-1} \end{array} \right) \quad \text{and} \quad x \rightarrow \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix}.$$

Then

$$\chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1} a_{n-1} = b.$$

Definition 5.9 Let $\{a_0, \ldots, a_{n-1}\} \subset \mathbb{R}^m$ and $\{\chi_0, \ldots, \chi_{n-1}\} \subset \mathbb{R}$. Then $\chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1} a_{n-1}$ is said to be a *linear combination* of the vectors $\{a_0, \ldots, a_{n-1}\}$.

We note that $Ax = b$ can be solved if and only if b equals a linear combination of the vectors that are the columns of A, by the definition of matrix-vector multiplication. This observation answers the question "Given a matrix A, for what right-hand side vector, b, does $Ax = b$ have a solution?" The answer is that there is a solution if and only if b is a linear combination of the columns (column vectors) of A.

Definition 5.10 The column space of $A \in \mathbb{R}^{m \times n}$ is the set of all vectors $b \in \mathbb{R}^m$ for which there exists a vector $x \in \mathbb{R}^n$ such that $Ax = b$.

Theorem 5.11 The column space of $A \in \mathbb{R}^{m \times n}$ is a subspace (of $\mathbb{R}^m$).

Proof: We need to show that the column space of A is closed under addition and scalar multiplication:

- Let $b_0, b_1 \in \mathbb{R}^m$ be in the column space of A. Then there exist $x_0, x_1 \in \mathbb{R}^n$ such that $Ax_0 = b_0$ and $Ax_1 = b_1$. But then $A(x_0 + x_1) = Ax_0 + Ax_1 = b_0 + b_1$ and thus $b_0 + b_1$ is in the column space of A.
- Let b be in the column space of A and $\alpha \in \mathbb{R}$. Then there exists a vector x such that $Ax = b$ and hence $\alpha Ax = \alpha b$. Since $A(\alpha x) = \alpha Ax = \alpha b$ we conclude that αb is in the column space of A.

Hence the column space of A is a subspace (of $\mathbb{R}^m$).

Example 5.12 Consider again

$$A = \begin{pmatrix} 3 & -1 & 2 \\ 1 & 2 & 0 \\ 4 & 1 & 2 \end{pmatrix}, \quad b_0 = \begin{pmatrix} 8 \\ -1 \\ 7 \end{pmatrix}.$$

Set this up as two appended systems, one for solving $Ax = b_0$ and the other for solving $Ax = 0$ (this will allow us to compare and contrast, which will lead to an interesting observation later on):

$$\left(\begin{array}{rrr|r} 3 & -1 & 2 & 8 \\ 1 & 2 & 0 & -1 \\ 4 & 1 & 2 & 7 \end{array}\right) \qquad \left(\begin{array}{rrr|r} 3 & -1 & 2 & 0 \\ 1 & 2 & 0 & 0 \\ 4 & 1 & 2 & 0 \end{array}\right). \tag{5.1}$$

Now, apply Gauss-Jordan elimination.

- It becomes convenient to swap the first and second equation:

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & -1 \\ 3 & -1 & 2 & 8 \\ 4 & 1 & 2 & 7 \end{array}\right) \qquad \left(\begin{array}{rrr|r} 1 & 2 & 0 & 0 \\ 3 & -1 & 2 & 0 \\ 4 & 1 & 2 & 0 \end{array}\right).$$

- Use the first row to eliminate the coefficients in the first column below the diagonal:

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & -1 \\ 0 & -7 & 2 & 11 \\ 0 & -7 & 2 & 11 \end{array}\right) \qquad \left(\begin{array}{rrr|r} 1 & 2 & 0 & 0 \\ 0 & -7 & 2 & 0 \\ 0 & -7 & 2 & 0 \end{array}\right).$$

- Use the second row to eliminate the coefficients in the second column below the diagonal:

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & -1 \\ 0 & -7 & 2 & 11 \\ 0 & 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{rrr|r} 1 & 2 & 0 & 0 \\ 0 & -7 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

- Divide the first and second row by the diagonal element:

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & -1 \\ 0 & 1 & -2/7 & -11/7 \\ 0 & 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{rrr|r} 1 & 2 & 0 & 0 \\ 0 & 1 & -2/7 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

- Use the second row to eliminate the coefficients in the second column above the diagonal:

$$\left(\begin{array}{rrr|r} 1 & 0 & 4/7 & 15/7 \\ 0 & 1 & -2/7 & -11/7 \\ 0 & 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{rrr|r} 1 & 0 & 4/7 & 0 \\ 0 & 1 & -2/7 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right). \tag{5.2}$$

Now, what does this mean? For now, we will focus only on the results for the appended system $\left(\begin{array}{c|c} A & b_0 \end{array}\right)$ on the left.

- We notice that $0 \times \chi_2 = 0$. So, there is no constraint on variable χ_2. As a result, we will call χ_2 a *free variable*.
- We see from the second row that $\chi_1 - 2/7\chi_2 = -11/7$ or $\chi_1 = -11/7 + 2/7\chi_2$. Thus, the value of χ_1 is constrained by the value given to χ_2.
- Finally, $\chi_0 + 4/7\chi_2 = 15/7$ or $\chi_0 = 15/7 - 4/7\chi_2$. Thus, the value of χ_0 is also constrained by the value given to χ_2.

We conclude that any vector of the form

$$\left(\begin{array}{r} 15/7 - 4/7\chi_2 \\ -11/7 + 2/7\chi_2 \\ \chi_2 \end{array}\right)$$

solves the linear system. We can rewrite this as

$$\left(\begin{array}{r} 15/7 \\ -11/7 \\ 0 \end{array}\right) + \chi_2 \left(\begin{array}{r} -4/7 \\ 2/7 \\ 1 \end{array}\right). \tag{5.3}$$

So, for each choice of χ_2, we get a solution to the linear system by plugging it into Equation (5.3).

Example 5.13 We now give a slightly "slicker" way to view Example 5.12. Consider again Equation (5.2):

$$\left(\begin{array}{ccr|r} 1 & 0 & 4/7 & 15/7 \\ 0 & 1 & -2/7 & -11/7 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

This represents

$$\left(\begin{array}{cc|r} 1 & 0 & 4/7 \\ 0 & 1 & -2/7 \\ \hline 0 & 0 & 0 \end{array}\right)\left(\begin{array}{c} \chi_0 \\ \chi_1 \\ \hline \chi_2 \end{array}\right) = \left(\begin{array}{r} 15/7 \\ -11/7 \\ 0 \end{array}\right).$$

Using blocked matrix-vector multiplication, we find that

$$\left(\begin{array}{c} \chi_0 \\ \chi_1 \end{array}\right) + \chi_2 \left(\begin{array}{r} 4/7 \\ -2/7 \end{array}\right) = \left(\begin{array}{r} 15/7 \\ -11/7 \end{array}\right)$$

and hence

$$\left(\begin{array}{c} \chi_0 \\ \chi_1 \end{array}\right) = \left(\begin{array}{r} 15/7 \\ -11/7 \end{array}\right) - \chi_2 \left(\begin{array}{r} 4/7 \\ -2/7 \end{array}\right)$$

which we can then turn into

$$\left(\begin{array}{c} \chi_0 \\ \chi_1 \\ \chi_2 \end{array}\right) = \left(\begin{array}{r} 15/7 \\ -11/7 \\ 0 \end{array}\right) + \chi_2 \left(\begin{array}{r} -4/7 \\ 2/7 \\ 1 \end{array}\right)$$

In the above example, we notice the following:

- Let $x_p = \left(\begin{array}{r} 15/7 \\ -11/7 \\ 0 \end{array}\right)$. Then $Ax_p = \left(\begin{array}{rrr} 3 & -1 & 2 \\ 1 & 2 & 0 \\ 4 & 1 & 2 \end{array}\right)\left(\begin{array}{r} 15/7 \\ -11/7 \\ 0 \end{array}\right) = \left(\begin{array}{r} 8 \\ -1 \\ 7 \end{array}\right)$. In other words, x_p is a *particular* solution to $Ax = b_0$. (Hence the p in the x_p.)
- Let $x_n = \left(\begin{array}{r} -4/7 \\ 2/7 \\ 1 \end{array}\right)$. Then $Ax_n = \left(\begin{array}{rrr} 3 & -1 & 2 \\ 1 & 2 & 0 \\ 4 & 1 & 2 \end{array}\right)\left(\begin{array}{r} -4/7 \\ 2/7 \\ 1 \end{array}\right) = \left(\begin{array}{r} 0 \\ 0 \\ 0 \end{array}\right)$. We will see that x_n is in the null space of A, to be defined later. (Hence the n in x_n.)
- Now, notice that for *any* α, $x_p + \alpha x_n$ is a solution to $Ax = b_0$:

$$A(x_p + \alpha x_n) = Ax_p + A(\alpha x_n) = Ax_p + \alpha A x_n = b_0 + \alpha \times 0 = b_0.$$

So, the system $Ax = b_0$ has many solutions (indeed, an infinite number of solutions). To characterize all solutions, it suffices to find one (nonunique) particular solution x_p that satisfies $Ax_p = b_0$. Now, for any vector x_n that has the property that $Ax_n = 0$, we know that $x_p + x_n$ is also a solution.

Definition 5.14 Let $A \in \mathbb{R}^{m \times n}$. Then the set of all vectors $x \in \mathbb{R}^n$ that have the property that $Ax = 0$ is called the *null space* of A and is denoted by $\mathcal{N}(A)$.

Theorem 5.15 The null space of $A \in \mathbb{R}^{m \times n}$ is a subspace of $\mathbb{R}^n$.

Proof: Clearly $\mathcal{N}(A)$ is a subset of $\mathbb{R}^n$. Now, assume that $x, y \in \mathcal{N}(A)$ and $\alpha \in \mathbb{R}$. Then $A(x + y) = Ax + Ay = 0$ and therefore $(x + y) \in \mathcal{N}(A)$. Also $A(\alpha x) = \alpha Ax = \alpha \times 0 = 0$ and therefore $\alpha x \in \mathcal{N}(A)$. Hence, $\mathcal{N}(A)$ is a subspace.

Notice that the zero vector (of appropriate length) is always in the null space of a matrix A.

Example 5.16 Let us use the last example, but with $Ax = b_1$: Let us set this up as an appended system

$$\left(\begin{array}{rrr|r} 3 & -1 & 2 & 5 \\ 1 & 2 & 0 & -1 \\ 4 & 1 & 2 & 7 \end{array}\right).$$

Now, apply Gauss-Jordan elimination.

- It becomes convenient to swap the first and second equation:

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & -1 \\ 3 & -1 & 2 & 5 \\ 4 & 1 & 2 & 7 \end{array}\right).$$

- Use the first row to eliminate the coefficients in the first column below the diagonal:

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & -1 \\ 0 & -7 & 2 & 8 \\ 0 & -7 & 2 & 11 \end{array}\right).$$

- Use the second row to eliminate the coefficients in the second column below the diagonal:

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & -1 \\ 0 & -7 & 2 & 8 \\ 0 & 0 & 0 & 3 \end{array}\right).$$

Now, what does this mean?

- We notice that $0 \times \chi_2 = 3$. This is a contradiction, and therefore this linear system has no solution!

Consider where we started: $Ax = b_1$ represents

$$\begin{array}{rrrcl} 3\chi_0 + & (-1)\chi_1 + & 2\chi_2 = & = & 5 \\ 1\chi_0 + & 2\chi_1 + & (0)\chi_2 = & = & -1 \\ 4\chi_0 + & \chi_1 + & 2\chi_2 = & = & 7. \end{array}$$

Now, the last equation is a linear combination of the first two. Indeed, add the first equation to the second, you get the third. Well, not quite: The last equation is actually inconsistent, because if you subtract the first two rows from the last, you don't get $0 = 0$. As a result, there is no way of simultaneously satisfying these equations.

5.2.1 A systematic procedure (first try)

Let us analyze what it is that we did in Examples 5.12 and 5.13.

- We start with $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$ where $Ax = b$ and x is to be computed.
- Append the system $\left(\begin{array}{c|c} A & b \end{array} \right)$.
- Use the Gauss-Jordan method to transform this appended system into the form

$$\left(\begin{array}{cc|c} I_{k\times k} & \tilde{A}_{TR} & \tilde{b}_T \\ 0_{(m-k)\times k} & 0_{(m-k)\times(n-k)} & \tilde{b}_B \end{array} \right), \quad (5.4)$$

  where $I_{k\times k}$ is the $k \times k$ identity matrix, $\tilde{A}_{TR} \in \mathbb{R}^{k\times(n-k)}$, $\tilde{b}_T \in \mathbb{R}^k$, and $\tilde{b}_B \in \mathbb{R}^{m-k}$.
- Now, if $\tilde{b}_B \neq 0$, then there is no solution to the system and we are done.
- Notice that (5.4) means that

$$\left(\begin{array}{cc} I_{k\times k} & \tilde{A}_{TR} \\ 0 & 0 \end{array} \right) \left(\begin{array}{c} x_T \\ x_B \end{array} \right) = \left(\begin{array}{c} \tilde{b}_T \\ 0 \end{array} \right).$$

  Thus, $x_T + \tilde{A}_{TR} x_B = \tilde{b}_T$. This translates to $x_T = \tilde{b}_T - \tilde{A}_{TR} x_B$ or

$$\left(\begin{array}{c} x_T \\ x_B \end{array} \right) = \left(\begin{array}{c} \tilde{b}_T \\ 0 \end{array} \right) + \left(\begin{array}{c} -\tilde{A}_{TR} \\ I_{(m-k)\times(m-k)} \end{array} \right) x_B.$$

- By taking $x_B = 0$, we find a *particular solution* $x_p = \begin{pmatrix} \tilde{b}_T \\ 0 \end{pmatrix}$.
- By taking, successively, $x_B = e_i$, $i = 0, \ldots, (m-k)-1$, we find vectors in the null space:
$$x_{n_i} = \begin{pmatrix} -\tilde{A}_{TR} \\ I_{(m-k)\times(m-k)} \end{pmatrix} e_i.$$
- The general solution is then given by
$$x_p + \alpha_0 x_{n_0} + \cdots + \alpha_{(m-k)-1} x_{n_{(m-k)-1}}.$$

Example 5.17 We now show how to use these insights to systematically solve the problem in Example 5.12. As in that example, create the appended systems for solving $Ax = b_0$ and $Ax = 0$ (Equation (5.1)).
$$\left(\begin{array}{rrr|r} 3 & -1 & 2 & 8 \\ 1 & 2 & 0 & -1 \\ 4 & 1 & 2 & 7 \end{array}\right) \qquad \left(\begin{array}{rrr|r} 3 & -1 & 2 & 0 \\ 1 & 2 & 0 & 0 \\ 4 & 1 & 2 & 0 \end{array}\right). \tag{5.5}$$
We notice that for $Ax = 0$ (the appended system on the right), the right-hand side never changes. It is always equal to zero. So, we don't really need to carry out all the steps for it, because everything to the left of the $|$ remains the same as it does for solving $Ax = b$. Carrying through with the Gauss-Jordan method, we end up with Equation (5.2):
$$\left(\begin{array}{rrr|r} 1 & 0 & 4/7 & 15/7 \\ 0 & 1 & -2/7 & -11/7 \\ 0 & 0 & 0 & 0 \end{array}\right)$$
Now, our procedure tells us that $x_p = \begin{pmatrix} 15/7 \\ -11/7 \\ 0 \end{pmatrix}$ is a particular solution: it solves $Ax = b$. Next, we notice that $\begin{pmatrix} \tilde{A}_{TR} \\ I_{(m-k)\times(m-k)} \end{pmatrix} = \begin{pmatrix} 4/7 \\ -2/7 \\ 1 \end{pmatrix}$ so that $A_{TR} = \begin{pmatrix} 4/7 \\ -2/7 \end{pmatrix}$, and $I_{(m-k)\times(m-k)} = 1$ (since there is only one free variable). So,
$$x_n = \begin{pmatrix} -\tilde{A}_{TR} \\ I_{(m-k)\times(m-k)} \end{pmatrix} e_0 = \begin{pmatrix} -4/7 \\ 2/7 \\ 1 \end{pmatrix} 1 = \begin{pmatrix} -4/7 \\ 2/7 \\ 1 \end{pmatrix}$$
The general solution is then given by
$$x = x_p + \alpha x_n = \begin{pmatrix} 15/7 \\ -11/7 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -4/7 \\ 2/7 \\ 1 \end{pmatrix},$$
for any scalar α.

5.2.2 A systematic procedure (second try)

Example 5.18 We now give an example where the procedure breaks down. Note: this example is borrowed from the book.

Consider $Ax = b$ where

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 2 \\ 10 \\ 10 \end{pmatrix}.$$

Let us apply Gauss-Jordan to this:

- Append the system:

$$\left(\begin{array}{rrrr|r} \mathbf{1} & 3 & 3 & 2 & 2 \\ 2 & 6 & 9 & 7 & 10 \\ -1 & -3 & 3 & 4 & 10 \end{array}\right).$$

- The boldfaced "1" is the pivot, in the first column. Subtract 2/1 times the first row and $(-1)/1$ times the first row from the second and third row, respectively:

$$\left(\begin{array}{rrrr|r} \mathbf{1} & 3 & 3 & 2 & 2 \\ 0 & 0 & \mathbf{3} & 3 & 6 \\ 0 & 0 & 6 & 6 & 12 \end{array}\right).$$

- The problem is that there is now a "0" in the second column. So, we skip it, and move on to the next column. The boldfaced "3" now becomes the pivot. Subtract 6/3 times the second row from the third row:

$$\left(\begin{array}{rrrr|r} \mathbf{1} & 3 & 3 & 2 & 2 \\ 0 & 0 & \mathbf{3} & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

- Divide each (nonzero) row by the pivot in that row to obtain

$$\left(\begin{array}{rrrr|r} \mathbf{1} & 3 & 3 & 2 & 2 \\ 0 & 0 & \mathbf{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

- We can only eliminate elements in the matrix above pivots. So, take 3/1 times the second row and subtract from the first row:

$$\left(\begin{array}{cccc|c} \mathbf{1} & 3 & 0 & -1 & -4 \\ 0 & 0 & \mathbf{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right). \tag{5.6}$$

- This does not have the form advocated in Equation 5.4. So, we remind ourselves of the fact that Equation 5.6 stands for

$$\left(\begin{array}{cccc} \mathbf{1} & 3 & 0 & -1 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \left(\begin{array}{c} \chi_{\mathbf{0}} \\ \chi_1 \\ \chi_{\mathbf{2}} \\ \chi_3 \end{array}\right) = \left(\begin{array}{c} -4 \\ 2 \\ 0 \end{array}\right) \tag{5.7}$$

  Notice that we can swap the second and third column of the matrix as long as we also swap the second and third element of the solution vector:

$$\left(\begin{array}{cc|cc} \mathbf{1} & 0 & 3 & -1 \\ 0 & \mathbf{1} & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \end{array}\right) \left(\begin{array}{c} \chi_{\mathbf{0}} \\ \chi_{\mathbf{2}} \\ \hline \chi_1 \\ \chi_3 \end{array}\right) = \left(\begin{array}{c} -4 \\ 2 \\ 0 \end{array}\right). \tag{5.8}$$

- Now, we notice that χ_1 and χ_3 are the free variables, and with those we can find equations for χ_0 and χ_2 as before.

- Also, we can now find vectors in the null space just as before. We just have to pay attention to the order of the unknowns (the order of the elements in the vector x).

In other words, a specific solution is now given by

$$x_p \left(\begin{array}{c} \chi_{\mathbf{0}} \\ \chi_{\mathbf{2}} \\ \hline \chi_1 \\ \chi_3 \end{array}\right) = \left(\begin{array}{c} -4 \\ 2 \\ 0 \\ 0 \end{array}\right)$$

and two linearly independent vectors in the null space are given by the columns of

$$\left(\begin{array}{cc} -3 & 1 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{array}\right)$$

giving us a general solution of

$$\left(\begin{array}{c} \chi_{\mathbf{0}} \\ \chi_{\mathbf{2}} \\ \hline \chi_1 \\ \chi_3 \end{array}\right) = \left(\begin{array}{c} -4 \\ 2 \\ 0 \\ 0 \end{array}\right) + \alpha \left(\begin{array}{c} 3 \\ 0 \\ 1 \\ 0 \end{array}\right) + \beta \left(\begin{array}{c} 1 \\ -1 \\ 0 \\ 1 \end{array}\right).$$

But notice that the order of the elements in the vector must be fixed (permuted):

$$\left(\begin{array}{c}\chi_0\\ \chi_1\\ \hline \chi_2\\ \chi_3\end{array}\right) = \left(\begin{array}{r}-4\\ 0\\ 2\\ 0\end{array}\right) + \alpha\left(\begin{array}{r}3\\ 1\\ 0\\ 0\end{array}\right) + \beta\left(\begin{array}{r}1\\ 0\\ -1\\ 1\end{array}\right).$$

Exercise 5.19 Let $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^m$, and $b \in \mathbb{R}^n$. Let $P \in \mathbb{R}^{m\times m}$ be a permutation matrix. Show that $AP^TPx = b$. Argue how this relates to the transition from Equation 5.7 to Equation 5.8.

Exercise 5.20 Complete Example 5.18 by computing a particular solution and two vectors in the null space (one corresponding to $\chi_1 = 1$, $\chi_3 = 0$ and the other to $\chi_1 = 0$, $\chi_3 = 1$).

5.3 Linear Independence

Definition 5.21 Let $\{a_0, \ldots, a_{n-1}\} \subset \mathbb{R}^m$. Then this set of vectors is said to be *linearly independent* if $\chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1}a_{n-1} = 0$ implies that $\chi_0 = \cdots = \chi_{n-1} = 0$. A set of vectors that is not linearly independent is said to be *linearly dependent*.

Notice that *if*

$$\chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1}a_{n-1} = 0 \quad \text{and} \quad \chi_j \neq 0,$$

then

$$\chi_j a_j = -\chi_0 a_0 + -\chi_1 a_1 - \cdots - \chi_{j-1}a_{j-1} - \chi_{j+1}a_{j+1} - \cdots - \chi_{n-1}a_{n-1}$$

and therefore

$$a_j = -\frac{\chi_0}{\chi_j}a_0 + -\frac{\chi_1}{\chi_j}a_1 - \cdots - \frac{\chi_{j-1}}{\chi_j}a_{j-1} - \frac{\chi_{j+1}}{\chi_j}a_{j+1} - \cdots - \frac{\chi_{n-1}}{\chi_j}a_{n-1}.$$

In other words, a_j can be written as a linear combination of the other $n-1$ vectors. This motivates the term *linearly independent* in the definition: none of the vectors can be written as a linear combination of the other vectors.

Theorem 5.22 Let $\{a_0, \ldots, a_{n-1}\} \subset \mathbb{R}^m$ and let $A = \left(\begin{array}{c|c|c} a_0 & \cdots & a_{n-1}\end{array}\right)$. Then the vectors $\{a_0, \ldots, a_{n-1}\}$ are linearly independent if and only if $\mathcal{N}(A) = \{0\}$.

Proof:

($\Rightarrow$) Assume $\{a_0, \ldots, a_{n-1}\}$ are linearly independent. We need to show that $\mathcal{N}(A) = \{0\}$. Assume $x \in \mathcal{N}(A)$. Then $Ax = 0$ implies that $0 = Ax = \left(\begin{array}{c|c|c} a_0 & \cdots & a_{n-1} \end{array} \right) \begin{pmatrix} \chi_0 \\ \vdots \\ \chi_{n-1} \end{pmatrix} =$ $\chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1} a_{n-1}$ and hence $\chi_0 = \cdots = \chi_{n-1} = 0$. Hence $x = 0$.

($\Leftarrow$) Notice that we are trying to prove $P \leftarrow Q$, where P represents "the vectors $\{a_0, \ldots, a_{n-1}\}$ are linearly independent" and Q represents "$\mathcal{N}(A) = \{0\}$". It suffices to prove the converse: $\neg P \rightarrow \neg Q$. Assume that $\{a_0, \ldots, a_{n-1}\}$ are *not* linearly independent. Then there exist $\{\chi_0, \cdots, \chi_{n-1}\}$ with at least one $\chi_j \neq 0$ such that $\chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1} a_{n-1} = 0$. Let $x = (\chi_0, \ldots, \chi_{n-1})^T$. Then $Ax = 0$ which means $x \in \mathcal{N}(A)$ and hence $\mathcal{N}(A) \neq \{0\}$.

Example 5.23 The columns of an identity matrix $I \in \mathbb{R}^{n \times n}$ form a linearly independent set of vectors.

Proof: Since I has an inverse (I itself) we know that $\mathcal{N}(I) = \{0\}$. Thus, by Theorem 5.22, the columns of I are linearly independent.

Example 5.24 The columns of $L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix}$ are linearly independent. If we consider

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} \chi_0 \\ \chi_1 \\ \chi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and simply solve this, we find that $\chi_0 = 0/1 = 0$, $\chi_1 = (0 - 2\chi_0)/(-1) = 0$, and $\chi_2 = (0 - \chi_0 - 2\chi_1)/(3) = 0$. Hence, $\mathcal{N}(L) = \{0\}$ (the zero vector) and we conclude, by Theorem 5.22, that the columns of L are linearly independent.

The last example motivates the following theorem:

Theorem 5.25 Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular matrix with nonzeroes on its diagonal. Then its columns are linearly independent.

Proof: Let L be as indicated and consider $Lx = 0$. If one solves this via whatever method one pleases, the solution $x = 0$ will emerge as the only solution. Thus $\mathcal{N}(L) = \{0\}$ and by Theorem 5.22, the columns of L are linearly independent.

Exercise 5.26 Let $U \in \mathbb{R}^{n\times n}$ be an upper triangular matrix with nonzeroes on its diagonal. Then its columns are linearly independent.

Exercise 5.27 Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular matrix with nonzeroes on its diagonal. Then its rows are linearly independent. (Hint: How do the rows of L relate to the columns of L^T?)

Example 5.28 The columns of $L = \left(\begin{array}{rrr} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & 2 & 3 \\ \hline -1 & 0 & -2 \end{array}\right)$ are linearly independent. If we consider

$$\left(\begin{array}{rrr} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & 2 & 3 \\ \hline -1 & 0 & -2 \end{array}\right) \begin{pmatrix} \chi_0 \\ \chi_1 \\ \chi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and simply solve this, we find that $\chi_0 = 0/1 = 0$, $\chi_1 = (0 - 2\chi_0)/(-1) = 0$, $\chi_2 = (0 - \chi_0 - 2\chi_1)/(3) = 0$. Hence, $\mathcal{N}(L) = \{0\}$ (the zero vector) and we conclude, by Theorem 5.22, that the columns of L are linearly independent.

This example motivates the following general observation:

Theorem 5.29 Let $A \in \mathbb{R}^{m\times n}$ have linearly independent columns and let $B \in \mathbb{R}^{k\times n}$. Then the matrices $\left(\begin{array}{c} A \\ \hline B \end{array}\right)$ and $\left(\begin{array}{c} B \\ \hline A \end{array}\right)$ have linearly independent columns.

Proof: Proof by contradiction. Assume that $\left(\frac{A}{B}\right)$ is not linearly independent. Then, by Theorem 5.22, there exists $x \in \mathbb{R}^n$ such that $x \neq 0$ and $\left(\frac{A}{B}\right) x = 0$. But that means that $\left(\frac{Ax}{Bx}\right) = \left(\frac{0}{0}\right)$, which in turn implies that $Ax = 0$. This contradicts the fact that the columns of A are linearly independent.

Corollary 5.30 Let $A \in \mathbb{R}^{m\times n}$. Then the matrix $\left(\frac{A}{I_{n\times n}}\right)$ has linearly independent columns.

Next, we observe that if one has a set of more than m vectors in $\mathbb{R}^m$, then they must be linearly dependent:

Theorem 5.31 Let $\{a_0, a_1, \vdots, a_{n-1}\} \in \mathbb{R}^m$ and $n > m$. Then these vectors are linearly dependent.

Proof: This proof is a bit more informal than I would like it to be: Consider the matrix $A = \left(a_0 \mid \cdots \mid a_{n-1} \right)$. If one apply the Gauss-Jordan method to this, at most m columns with pivots will be encountered. The other $n - m$ columns correspond to free variables, which allow us to construct nonzero vectors x so that $Ax = 0$.

5.4 Bases

Definition 5.32 Let $\{v_0, v_1, \cdots, v_{n-1}\} \subset \mathbb{R}^m$. Then the *span* of these vectors, $\text{Span}(\{v_0, v_1, \cdots, v_{n-1}\})$, is said to be the space of all vectors that are a linear combination of this set of vectors.

Notice that $\text{Span}(\{v_0, v_1, \cdots, v_{n-1}\})$ equals the column space of the matrix $\left(v_0 \mid \cdots \mid v_n \right)$.

Definition 5.33 Let $\mathbf{V}$ be a subspace of $\mathbb{R}^m$. Then the set $\{v_0, v_1, \cdots, v_{n-1}\} \subset \mathbb{R}^m$ is said to be a *spanning set* for $\mathbf{V}$ if $\text{Span}(\{v_0, v_1, \cdots, v_{n-1}\}) = \mathbf{V}$.

Definition 5.34 Let $\mathbf{V}$ be a subspace of $\mathbb{R}^m$. Then the set $\{v_0, v_1, \cdots, v_{n-1}\} \subset \mathbb{R}^m$ is said to be a *basis* for $\mathbf{V}$ if (1) $\{v_0, v_1, \cdots, v_{n-1}\}$ are linearly independent and (2) $\text{Span}(\{v_0, v_1, \cdots, v_{n-1}\}) = \mathbf{V}$.

The first condition says that there aren't more vectors than necessary in the set. The second says there are enough to be able to generate $\mathbf{V}$.

Example 5.35 The vectors $\{e_0, \ldots, e_{m-1}\} \subset \mathbb{R}^m$, where e_j equals the jth column of the identity, are a basis for $\mathbb{R}^m$.

Note: these vectors are linearly independent and any vector $x \in \mathbb{R}^m$ with $x = \begin{pmatrix} \chi_0 \\ \vdots \\ \chi_{m-1} \end{pmatrix}$ can be written as the linear combination $\chi_0 e_0 + \cdots + \chi_{m-1} e_{m-1}$.

Example 5.36 Let $\{a_0, \ldots, a_{m-1}\} \subset \mathbb{R}^m$ and let $A = \left(a_0 \mid \cdots \mid a_{m-1} \right)$ be invertible. Then $\{a_0, \ldots, a_{m-1}\} \subset \mathbb{R}^m$ form a basis for $\mathbb{R}^m$.

Note: The fact that A is invertible means there exists A^{-1} such that $A^{-1}A = I$. Since $Ax = 0$ means $x = A^{-1}Ax = A^{-1}0 = 0$, the columns of A are linearly independent. Also, given any vector $y \in \mathbb{R}^m$, there exists a vector $x \in \mathbb{R}^m$ such that $Ax = y$ (namely $x = A^{-1}y$). Letting $x = \begin{pmatrix} \chi_0 \\ \vdots \\ \chi_{m-1} \end{pmatrix}$ we find that $y = \chi_0 a_0 + \cdots + \chi_{m-1} a_{m-1}$ and hence every vector in $\mathbb{R}^m$ is a linear combination of the set $\{a_0, \ldots, a_{m-1}\} \subset \mathbb{R}^m$.

Now here comes a very important insight:

Theorem 5.37 Let $\mathbf{V}$ be a subspace of $\mathbb{R}^m$ and let $\{v_0, v_1, \cdots, v_{n-1}\} \subset \mathbb{R}^m$ and $\{w_0, w_1, \cdots, w_{k-1}\} \subset \mathbb{R}^m$ both be bases for $\mathbf{V}$. Then $k = n$. In other words, the number of vectors in a basis is unique.

Proof: Proof by contradiction. Without loss of generality, let us assume that $k > n$. (Otherwise, we can switch the roles of the two sets.) Let $V = \left(v_0 \mid \cdots \mid v_{n-1} \right)$ and $W = \left(w_0 \mid \cdots \mid w_{k-1} \right)$. Let x_j have the property that $w_j = V x_j$. (We know such a vector x_j exists because V spans $\mathbf{V}$ and $w_j \in \mathbf{V}$.) Then $W = VX$, where $X = \left(x_0 \mid \cdots \mid x_{k-1} \right)$. Now, $X \in \mathbb{R}^{n \times k}$ and recall that $k > n$. This means that $\mathcal{N}(X)$ contains nonzero vectors (why?). Let $y \in \mathcal{N}(X)$. Then $Wy = VXy = V(Xy) = V(0) = 0$, which contradicts the fact

---

that $\{w_0, w_1, \cdots, w_{k-1}\}$ are linearly independent, and hence this set cannot be a basis for $\mathbf{V}$.

Note: generally speaking, there are an infinite number of bases for a given subspace. (The exception is the subspace $\{0\}$.) However, the number of vectors in each of these bases is always the same. This allows us to make the following definition:

Definition 5.38 The dimension of a subspace $\mathbf{V}$ equals the number of vectors in a basis for that subspace.

A basis for a subspace $\mathbf{V}$ can be derived from a spanning set of a subspace $\mathbf{V}$ by, one-to-one, removing vectors from the set that are dependent on other remaining vectors until the remaining set of vectors is linearly independent , as a consequence of the following observation:

Definition 5.39 Let $A \in \mathbb{R}^{m\times n}$. The *rank* of A equals the number of vectors in a basis for the column space of A. We will let $\text{rank}(A)$ denote that rank.

Theorem 5.40 Let $\{v_0, v_1, \cdots, v_{n-1}\} \subset \mathbb{R}^m$ be a spanning set for subspace $\mathbf{V}$ and assume that v_i equals a linear combination of the other vectors. Then $\{v_0, v_1, \cdots, v_{i-1}, v_i, v_{i-1}, \cdots, v_{n-1}\}$ is a spanning set of $\mathbf{V}$.

Similarly, a set of linearly independent vectors that are in a subspace $\mathbf{V}$ can be "built up" to be a basis by successively adding vectors that are in $\mathbf{V}$ to the set while maintaining that the vectors in the set remain linearly independent until the resulting is a basis for $\mathbf{V}$.

Theorem 5.41 Let $\{v_0, v_1, \cdots, v_{n-1}\} \subset \mathbb{R}^m$ be linearly independent and assume that $\{v_0, v_1, \cdots, v_{n-1}\} \subset V$. Then this set of vectors is either a spanning set for V or there exists $w \in \mathbf{V}$ such that $\{v_0, v_1, \cdots, v_{n-1}, w\}$ are linearly independent.

5.5 Exercises

(Most of these exercises are borrowed from "Linear Algebra and Its Application" by Gilbert Strang.)

1. Which of the following subsets of $\mathbb{R}^3$ are actually subspaces?

   (a) The plane of vectors $x = (\chi_0, \chi_1, \chi_2)^T \in \mathbb{R}^3$ such that the first component $\chi_0 = 0$. In other words, the set of all vectors

$$\begin{pmatrix} 0 \\ \chi_1 \\ \chi_2 \end{pmatrix}$$

where $\chi_1, \chi_2 \in \mathbb{R}$.

(b) The plane of vectors x with $\chi_0 = 1$.

(c) The vectors x with $\chi_1\chi_2 = 0$ (this is a union of two subspaces: those vectors with $\chi_1 = 0$ and those vectors with $\chi_2 = 0$).

(d) All combinations of two given vectors $(1,1,0)^T$ and $(2,0,1)^T$.

(e) The plane of vectors $x = (\chi_0, \chi_1, \chi_2)^T$ that satisfy $\chi_2 - \chi_1 + 3\chi_0 = 0$.

2. Describe the column space and nullspace of the matrices

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

3. Let $\mathbf{P} \subset \mathbb{R}^3$ be the plane with equation $x + 2y + z = 6$. What is the equation of the plane $\mathbf{P}_0$ through the origin parallel to $\mathbf{P}$? Are $\mathbf{P}$ and/or $\mathbf{P}_0$ subspaces of R^3?

4. Let $\mathbf{P} \subset \mathbb{R}^3$ be the plane with equation $x + y - 2z = 4$. Why is this not a subspace? Find two vectors, x and y, that are in $\mathbf{P}$ but with the property that $x + y$ is not.

5. Find the echolon form U, the free variables, and the special (particular) solution of $Ax = b$ for

(a) $A = \begin{pmatrix} 0 & 1 & 0 & 3 \\ 0 & 2 & 0 & 6 \end{pmatrix}$, $b = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$. When does $Ax = b$ have a solution? (When $\beta_1 = ?$.) Give the complete solution.

(b) $A = \begin{pmatrix} 0 & 0 \\ 1 & 2 \\ 0 & 0 \\ 3 & 6 \end{pmatrix}$, $b = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$. When does $Ax = b$ have a solution? (When ...) Give the complete solution.

6. Write the complete soluton $x = x_p + x_n$ to these systems:

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 5 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

7. Which of these statements is a correct definition of the rank of a given matrix $A \in \mathbb{R}^{m\times n}$? (Indicate all correct ones.)

(a) The number of nonzero rows in the row reduced form of A.

(b) The number of columns minus the number of rows, $n-m$.

(c) The number of columns minus the number of free columns in the row reduced form of A. (Note: a free column is a column that does not contain a pivot.)

(d) The number of 1s in the row reduced form of A.

8. Let $A, B \in \mathbb{R}^{m\times n}$, $u \in \mathbb{R}^m$, and $v \in \mathbb{R}^n$. The operation $B := A + \alpha u v^T$ is often called a rank-1 update. Why?

9. Find the complete solution of

$$\begin{array}{rrrcr} x & +3y & +z & = & 1 \\ 2x & +6y & +9z & = & 5 \\ -x & -3y & +3z & = & 5 \end{array}.$$

10. Find the complete solution of

$$\begin{pmatrix} 1 & 3 & 1 & 2 \\ 2 & 6 & 4 & 8 \\ 0 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}.$$

5.6 The Answer to Life, The Universe, and Everything

We complete this chapter by showing how many answers about subspaces can be answered from the upper-echolon form of the linear system.

To do so, consider again

$$\begin{pmatrix} 1 & 3 & 1 & 2 \\ 2 & 6 & 4 & 8 \\ 0 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$$

from Question 10 in the previous section. Reducing this to upper echolon format yields

$$\left(\begin{array}{cccc|c} 1 & 3 & 1 & 2 & 1 \\ 2 & 6 & 4 & 8 & 3 \\ 0 & 0 & 2 & 4 & 1 \end{array}\right) \quad \rightarrow \quad \left(\begin{array}{cccc|c} \boxed{1} & 3 & 1 & 2 & 1 \\ 0 & 0 & \boxed{2} & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Here the boxed entries are the pivots (the first nonzero entry in each row) and they identify that the corresponding variables (x and z) are dependent variables while the other variables (y and t) are free variables.

Give the general solution to the problem To find the general solution to problem, you recognize that there are two free variables (y and t) and a general solution can thus be given by

$$\begin{pmatrix} \Box \\ 0 \\ \Box \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} \Box \\ 1 \\ \Box \\ 0 \end{pmatrix} + \beta \begin{pmatrix} \Box \\ 0 \\ \Box \\ 1 \end{pmatrix}.$$

Here $x_p = \begin{pmatrix} \Box \\ 0 \\ \Box \\ 0 \end{pmatrix}$ is a particular (special) solution that solves the system. To obtain it, you set the free variables to zero and solve for the values in the boxes:

$$\begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ 0 \\ z \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

or

$$\begin{array}{rcl} x \quad +z & = & 1 \\ 2z & = & 1 \end{array}$$

so that $z = 1/2$ and $x = 1/2$ yielding a particular solution $x_p = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{pmatrix}$.

Next, we have to find the two vectors in the null space of the matrix. To obtain the first, we set the first free variable to one and the other(s) to zero:

$$\begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ 1 \\ z \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

or

$$\begin{array}{cccc} x & +3 \times 1 & +z & = & 0 \\ & & 2z & = & 0 \end{array}$$

so that $z = 0$ and $x = -3$, yielding the first vector in the null space $x_{n_0} = \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

To obtain the second, we set the second free variable to one and the other(s) to zero:

$$\begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ 0 \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

or

$$\begin{array}{ccccc} x & +z & +2 \times 1 & = & 0 \\ & 2z & +4 \times 1 & = & 0 \end{array}$$

so that $z = -4/2 = -2$ and $x = -z - 2 = 0$, yielding the second vector in the null space $x_{n_1} = \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}$.

And thus the general solution is given by

$$\begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix},$$

where α and β are scalars.

Find a specific (particular) solution to the problem The above procedure yields the particular solution as part of the first step.

Find vectors in the null space The first step is to figure out how many (linear independent) vectors there are in the null space. This equals the number of free variables. The above procedure then gives you a step-by-step procedure for finding that many linearly independent vectors in the null space.

Find linearly independent columns in the original matrix **Note: this is equivalent to finding a basis for the column space of the matrix.**

To find the linearly independent columns, you look at the upper echolon form of the matrix:

$$\begin{pmatrix} \boxed{1} & 3 & 1 & 2 \\ 0 & 0 & \boxed{2} & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

with the pivots highlighted. The columns that have pivots in them are linearly independent. The corresponding columns in the original matrix are linearly independent:

$$\left(\begin{array}{|c|c|c|c} \hline 1 & 3 & \multicolumn{1}{|c|}{1} & 2 \\ 2 & 6 & \multicolumn{1}{|c|}{4} & 8 \\ 0 & 0 & \multicolumn{1}{|c|}{2} & 4 \\ \hline \end{array}\right).$$

Thus, in our example, the answer is $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$ (the first and third column).

Dimension of the Column Space (Rank of the Matrix) The following are all equal:

- The dimension of the column space.
- The rank of the matrix.
- The number of dependent variables.
- The number of nonzero rows in the upper echelon form.
- The number of columns in the matrix minus the number of free variables.
- The number of columns in the matrix minus the dimension of the null space.
- The number of linearly independent columns in the matrix.
- The number of linearly independent rows in the matrix.

Find a basis for the row space of the matrix. The row space (we will see in the next chapter) is the space spanned by the rows of the matrix (viewed as column vectors). Reducing a matrix to upper echelon form merely takes linear combinations of the rows of the matrix. What this means is that the space spanned by the rows of the original matrix is the same space as is spanned by the rows of the matrix in upper echolon form. Thus, all you need to do is list the rows in the matrix in upper echolon form, as column vectors.

For our example this means a basis for the row space of the matrix is given by $\begin{pmatrix} 1 \\ 3 \\ 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix}$.