A.3: Elimination

Linear Systems of Equations

One application of matrices is to solve systems of linear equations. Consider the following system of linear equations.

\[
\begin{align}
2x_1 + 2x_2 + 2x_3 &= 2, \\
x_1 + x_2 + 3x_3 &= 5, \\
x_1 + 4x_2 + x_3 &= 10.
\end{align}
\]

There is a systematic procedure called elimination to solve such a system. In this procedure, we attempt to eliminate each variable from all but one equation. We want to end up with equations such as \(x_3 = 2\), where we can just read off the answer.

We write a system of linear equations as a matrix equation: \(A\vec{x} = \vec{b}\). The system \((eqref{eq:1})\) is written as:

\[
\begin{bmatrix}
2 & 2 & 2 \\
1 & 1 & 3 \\
1 & 4 & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix} =
\begin{bmatrix}
2 \\
5 \\
10
\end{bmatrix}.
\]

If we knew the inverse of \((A)\), then we would be done; we would simply solve the equation: \(\vec{x} = A^{-1}\vec{b}\). Well, but that is part of the problem, we do not know how to compute the inverse for matrices bigger than \((2 \times 2)\). We will see later that to compute the inverse we are really solving \((A\vec{x} = \vec{b})\) for several different \((\vec{b})\). In other words, we will need to do elimination to find \((A^{-1})\). In addition, we may wish to solve \((A\vec{x} = \vec{b})\) if \((A)\) is not invertible, or perhaps not even square.

Let us return to the equations themselves and see how we can manipulate them. There are a few operations we can perform on the equations that do not change the solution. First, perhaps an operation that may seem stupid, we can swap two equations in \((eqref{eq:1})\):

\[
\begin{align}
\phantom{9} x_1 + \phantom{9} x_2 + \phantom{9} x_3 & = 5 , \\
\phantom{9} x_1 + \phantom{9} x_2 + \phantom{9} x_3 & = 2 , \\
\phantom{9} x_1 + 4 x_2 + \phantom{9} x_3 & = 10 .
\end{align}
\]

UC Davis ChemWiki is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License.
Clearly these new equations have the same solutions \((x_1, x_2, x_3)\). A second operation is that we can multiply an equation by a nonzero number. For example, we multiply the third equation in \((\text{eq:1})\) by 3:
\[
\begin{align}
2 x_1 + 2 x_2 + 2 x_3 & = 2 , \\
x_1 + 3 x_2 + 3 x_3 & = 5 , \\
3 x_1 + 12 x_2 + 3 x_3 & = 30 .
\end{align}
\]

Finally, we can add a multiple of one equation to another equation. For instance, we add 3 times the third equation in \((\text{eq:1})\) to the second equation:
\[
\begin{align}
2 x_1 + 2 x_2 + 2 x_3 & = 2 , \\
(1+3) x_1 + (1+12) x_2 + (3+3) x_3 & = 5+30 , \\
x_1 + 4 x_2 + x_3 & = 10 .
\end{align}
\]
The same \((x_1, x_2, x_3)\) should still be solutions to the new equations. These were just examples; we did not get any closer to the solution. We must do these three operations in some more logical manner, but it turns out these three operations suffice to solve every linear equation.

The first thing is to write the equations in a more compact manner. Given \(A \vec{x} = \vec{b}\) we write down the so-called augmented matrix \([A:|:\vec{b}]\) where the vertical line is just a marker for us to know where the of the equation starts. For the system \((\text{eq:1})\) the augmented matrix is
\[
\left[\begin{array}{ccc|c}
2 & 2 & 2 & 2 \\
1 & 1 & 3 & 5 \\
1 & 4 & 1 & 10
\end{array}\right].
\]
The entire process of elimination, which we will describe, is often applied to any sort of matrix, not just an augmented matrix. Simply think of the matrix as the \((3 \times 4)\) matrix \(\begin{bmatrix} 2 & 2 & 2 & 2 \\
1 & 1 & 3 & 5 \\
1 & 4 & 1 & 10 \end{bmatrix}\).

**Echelon Form and Elementary Operations**

We apply the three operations above to the matrix. We call these the *elementary operations* or *elementary row operations*. Translating the operations to the matrix setting, the operations become:

i. Swap two rows.
ii. Multiply a row by a nonzero number.
iii. Add a multiple of one row to another row.

We run these operations until we get into a state where it is easy to read off the answer, or until we get into a contradiction indicating no solution.

More specifically, we run the operations until we obtain the so-called *row echelon form*. Let us call the first (from the left) nonzero entry in each row the leading entry. A matrix is in *row echelon form* if the following conditions are satisfied:

i. The leading entry in any row is strictly to the right of the leading entry of the row above.
ii. Any zero rows are below all the nonzero rows.
iii. All leading entries are \((1)\).

A matrix is in *reduced row echelon form* if furthermore the following condition is satisfied.

iv. All the entries above a leading entry are zero.

Note that the definition applies to matrices of any size.
Example \(\PageIndex{1}\)

The following matrices are in row echelon form. The leading entries are marked:
\[
\begin{bmatrix}
\fbox{1} & 2 & 9 & 3 \\
0 & \fbox{1} & 5 \\
0 & 0 & \fbox{1}
\end{bmatrix} \quad
\begin{bmatrix}
\fbox{1} & -1 & -3 \\
0 & \fbox{1} \\
0 & 0 & \fbox{1}
\end{bmatrix} \quad
\begin{bmatrix}
\fbox{1} & 2 & 1 \\
0 & \fbox{1} \\
0 & 0 & \fbox{1}
\end{bmatrix} \quad
\begin{bmatrix}
0 & \fbox{1} & -5 & 2 \\
0 & 0 & \fbox{1} \\
0 & 0 & 0 & \fbox{1}
\end{bmatrix}
\]
None of the matrices above are in reduced row echelon form. For example, in the first matrix none of the entries above the second and third leading entries are zero; they are 9, 3, and 5. The following matrices are in reduced row echelon form. The leading entries are marked:
\[
\begin{bmatrix}
\fbox{1} & 3 & 0 & 8 \\
0 & \fbox{1} & 6 \\
0 & 0 & 0 
\end{bmatrix} \quad
\begin{bmatrix}
\fbox{1} & 0 & 2 & 0 \\
0 & \fbox{1} & 3 \\
0 & 0 & 0 
\end{bmatrix} \quad
\begin{bmatrix}
\fbox{1} & 0 & 3 \\
0 & \fbox{1} & -2 \\
0 & 0 & 0 
\end{bmatrix} \quad
\begin{bmatrix}
0 & \fbox{1} & -5 & 2 \\
0 & 0 & \fbox{1} \\
0 & 0 & 0 & \fbox{1}
\end{bmatrix}
\]

The procedure we will describe to find a reduced row echelon form of a matrix is called \textit{Gauss-Jordan elimination}. The first part of it, which obtains a row echelon form, is called \textit{Gaussian elimination} or \textit{row reduction}. For some problems, a row echelon form is sufficient, and it is a bit less work to only do this first part.

To attain the row echelon form we work systematically. We go column by column, starting at the first column. We find topmost entry in the first column that is not zero, and we call it the \textit{pivot}. If there is no nonzero entry we move to the next column. We swap rows to put the row with the pivot as the first row. We divide the first row by the pivot to make the pivot entry be a 1. Now look at all the rows below and subtract the correct multiple of the pivot row so that all the entries below the pivot become zero.

After this procedure we forget that we had a first row (it is now fixed), and we forget about the column with the pivot and all the preceding zero columns. Below the pivot row, all the entries in these columns are just zero. Then we focus on the smaller matrix and we repeat the steps above.

It is best shown by example, so let us go back to the example from the beginning of the section. We keep the vertical line in the matrix, even though the procedure works on any matrix, not just an augmented matrix. We start with the first column and we locate the pivot, in this case the first entry of the first column.
\[
\begin{bmatrix}
\fbox{2} & 2 & 2 & 2 \\
1 & 1 & 3 & 5 \\
1 & 4 & 1 & 10
\end{bmatrix}
\]
We multiply the first row by \(\frac{1}{2}\).\[
\begin{bmatrix}
\fbox{1} & 1 & 1 & 1 \\
0 & 0 & 2 & 4 \\
0 & 3 & 0 & 9
\end{bmatrix}
\]
We are done with the first column and the first row for now. We almost pretend the matrix doesn’t have the first column and the first row. OK, look at the second column, and notice that now the pivot is in the third row. We swap rows. And we divide the pivot row by 3. We do not need to subtract anything as everything below the pivot is already zero. We move on, we again start ignoring the second row and second column and focus on...
The matrix is now in row echelon form. The equation corresponding to the last row is \( x_3 = 2 \). We know \( x_3 \) and we could substitute it into the first two equations to get equations for \( x_1 \) and \( x_2 \). Then we could do the same thing with \( x_2 \), until we solve for all 3 variables. This procedure is called backsubstitution and we can achieve it via elementary operations. We start from the lowest pivot (leading entry in the row echelon form) and subtract the right multiple from the row above to make all the entries above this pivot zero. Then we move to the next pivot and so on. After we are done, we will have a matrix in reduced row echelon form.

We continue our example. Subtract the last row from the first to get \( \begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array} \). The entry above the pivot in the second row is already zero. So we move onto the next pivot, the one in the second row. We subtract this row from the top row to get \( \begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array} \). The matrix is in reduced row echelon form.

If we now write down the equations for \( x_1, x_2, x_3 \), we find \( x_1 = -4, x_2 = 3, x_3 = 2 \). In other words, we have solved the system.

### Non-Unique Solutions and Inconsistent Systems

It is possible that the solution of a linear system of equations is not unique, or that no solution exists. Suppose for a moment that the row echelon form we found was \( \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array} \). Then we have an equation \( 0=1 \) coming from the last row. That is impossible and the equations are what we call inconsistent. There is no solution to \( A \vec{x} = \vec{b} \).

On the other hand, if we find a row echelon form \( \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array} \), then there is no issue with finding solutions. In fact, we will find way too many. Let us continue with backsubstitution (subtracting 3 times the third row from the first) to find the reduced row echelon form and let’s mark the pivots. \( \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array} \). The last row is all zeros; it just says \( 0=0 \) and we ignore it. The two remaining equations are \( x_1 + 2 x_2 = -5, x_3 = 3 \). Let us solve for the variables that corresponded to the pivots, that is \( x_1 \) and \( x_3 \) as there was a pivot in the first column and in the third column: \( \begin{array}{c} x_1 = -2 x_2 - 5 \\ x_3 = 3 \end{array} \). The variable \( x_2 \) can be anything you wish and we still get a solution. The \( x_2 \) is called a free variable. There are infinitely many solutions, one for every choice of \( x_2 \). If we pick \( x_2 = 0 \), then \( x_1 = -5 \), and \( x_3 = 3 \) give a solution. But we also get a solution by picking say \( x_2 = 1 \), in which case \( x_1 = -9 \) and \( x_3 = 3 \), or by picking \( x_2 = -5 \) in which case \( x_1 = 5 \) and \( x_3 = 3 \).

The general idea is that if any row has all zeros in the columns corresponding to the variables, but a nonzero entry in the
column corresponding to the right-hand side \( \langle \vec{b} \rangle \), then the system is inconsistent and has no solutions. In other words, the system is inconsistent if you find a pivot on the right side of the vertical line drawn in the augmented matrix. Otherwise, the system is consistent, and at least one solution exists.

Suppose the system is consistent (at least one solution exists):

i. If every column corresponding to a variable has a pivot element, then the solution is unique.

ii. If there are columns corresponding to variables with no pivot, then those are free variables that can be chosen arbitrarily, and there are infinitely many solutions.

When \( \langle \vec{b} \rangle = \langle \vec{0} \rangle \), we have a so-called homogeneous matrix equation \( \langle A \vec{x} = \vec{0} \rangle \). There is no need to write an augmented matrix in this case. As the elementary operations do not do anything to a zero column, it always stays a zero column. Moreover, \( \langle A \vec{x} = \vec{0} \rangle \) always has at least one solution, namely \( \langle \vec{x} = \vec{0} \rangle \). Such a system is always consistent. It may have other solutions: If you find any free variables, then you get infinitely many solutions.

The set of solutions of \( \langle A \vec{x} = \vec{0} \rangle \) comes up quite often so people give it a name. It is called the nullspace or the kernel of \( \langle A \rangle \). One place where the kernel comes up is invertibility of a square matrix \( \langle A \rangle \). If the kernel of \( \langle A \rangle \) contains a nonzero vector, then it contains infinitely many vectors (there was a free variable). But then it is impossible to invert \( \langle \vec{0} \rangle \), since infinitely many vectors go to \( \langle \vec{0} \rangle \), so there is no unique vector that \( \langle A \rangle \) takes to \( \langle \vec{0} \rangle \). So if the kernel is nontrivial, that is, if there are any nonzero vectors in the kernel, in other words, if there are any free variables, or in yet other words, if the row echelon form of \( \langle A \rangle \) has columns without pivots, then \( \langle A \rangle \) is not invertible. We will return to this idea later.

---

**Linear Independence and Rank**

If rows of a matrix correspond to equations, it may be good to find out how many equations we really need to find the same set of solutions. Similarly, if we find a number of solutions to a linear equation \( \langle A \vec{x} = \vec{0} \rangle \), we may ask if we found enough so that all other solutions can be formed out of the given set. The concept we want is that of linear independence. That same concept is useful for differential equations, for example in Chapter 2.

Given row or column vectors \( \langle \vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n \rangle \), a linear combination is an expression of the form \( \langle \alpha_1 \vec{y}_1 + \alpha_2 \vec{y}_2 + \cdots + \alpha_n \vec{y}_n \rangle \) where \( \alpha_1, \alpha_2, \ldots, \alpha_n \) are all scalars. For example, \( \langle 3 \vec{y}_1 + \vec{y}_2 - 5 \vec{y}_3 \rangle \) is a linear combination of \( \langle \vec{y}_1, \vec{y}_2, \vec{y}_3 \rangle \).

We have seen linear combinations before. The expression \( \langle A \vec{x} \rangle \) is a linear combination of the columns of \( \langle A \rangle \), while \( \langle A^T \vec{x} \rangle \) is a linear combination of the rows of \( \langle A \rangle \).

The way linear combinations come up in our study of differential equations is similar to the following computation. Suppose that \( \langle \vec{x}_1 \rangle, \langle \vec{x}_2 \rangle, \ldots, \langle \vec{x}_n \rangle \) are solutions to \( \langle A \vec{x}_1 = \vec{0} \rangle, \langle A \vec{x}_2 = \vec{0} \rangle, \ldots, \langle A \vec{x}_n = \vec{0} \rangle \). Then the linear combination \( \langle \vec{y} = \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n \rangle \) is also a solution.
\( \alpha_n \vec{x}_n \) is a solution to \( (A \vec{y} = \vec{0}) \): \[
\begin{align*}
A \vec{y} &= A (\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n) \\
&= \alpha_1 A \vec{x}_1 + \alpha_2 A \vec{x}_2 + \cdots + \alpha_n A \vec{x}_n = \alpha_1 \vec{0} + \alpha_2 \vec{0} + \cdots + \alpha_n \vec{0} = \vec{0}.
\end{align*}
\] So if you have found enough solutions, you have them all. The question is, when did we find enough of them?

We say the vectors \( \{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n\} \) are **linearly independent** if the only solution to \( \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n = \vec{0} \) is \( \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \). Otherwise, we say the vectors are **linearly dependent**.

For example, the vectors \( \begin{bmatrix} 1 \\ 2 \end{bmatrix} \) and \( \begin{bmatrix} 0 \\ 1 \end{bmatrix} \) are linearly independent. Let's try: \[
\alpha_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ 2 \alpha_1 + \alpha_2 \end{bmatrix} = \vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
\] So \( \alpha_1 = 0 \), and then it is clear that \( \alpha_2 = 0 \) as well. In other words, the two vectors are linearly independent.

If a set of vectors is linearly dependent, that is, some of the \( \alpha_j \)'s are nonzero, then we can solve for one vector in terms of the others. Suppose \( \alpha_1 \not= 0 \). Since \( \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n = \vec{0} \), then \( \vec{x}_1 = \frac{-\alpha_2}{\alpha_1} \vec{x}_2 - \frac{-\alpha_3}{\alpha_1} \vec{x}_3 + \cdots + \frac{-\alpha_n}{\alpha_1} \vec{x}_n \). For example, \[
\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - 2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ -9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.
\] So \( \alpha_1 = 0 \), and then it is clear that \( \alpha_2 = 0 \) as well. In other words, the two vectors are linearly independent.

Given a set of vectors, we may not be interested in just finding if they are linearly independent or not, we may be interested in finding a linearly independent subset. Or perhaps we may want to find some other vectors that give the same linear combinations and are linearly independent. The way to figure this out is to form a matrix out of our vectors. If we have row vectors we consider them as rows of a matrix. If we have column vectors we consider them columns of a matrix. The set of all linear combinations of a set of vectors is called their **span**. \( \text{\operatorname{span}} \{ \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n \} = \{ \text{Set of all linear combinations of } \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n \} \).

Given a matrix \( A \), the maximal number of linearly independent rows is called the **rank** of \( A \), and we write it \( \text{\operatorname{rank}} A \). The second and third row are multiples of the first one. We cannot choose more than one row and still have a linearly independent set. But what is \( \text{\operatorname{rank}} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \) ? That seems to be a tougher question to answer. The first two rows are linearly independent (neither is a multiple of the other), so the rank is at least two. If we would set up the equations for the \( \{\alpha_1, \alpha_2, \alpha_3\} \), we would find a system with infinitely many solutions. One solution is \( \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} - 2 \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \). So the set of all three rows is linearly dependent, the rank cannot be 3. Therefore the rank is 2.
But how can we do this in a more systematic way? We find the row echelon form! \[
\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{bmatrix}
\]
The elementary row operations do not change the set of linear combinations of the rows (that was one of the main reasons for defining them as they were). In other words, the span of the rows of the \((A)\) is the same as the span of the rows of the row echelon form of \((A)\). In particular, the number of linearly independent rows is the same. And in the row echelon form, all nonzero rows are linearly independent. This is not hard to see. Consider the two nonzero rows in the example above. Suppose we tried to solve for the \(\alpha_1\) and \(\alpha_2\) in \[
\begin{bmatrix} 1 & 2 & 3 \\
0 & 1 & 2
\end{bmatrix}
\]
Since the first column of the row echelon matrix has zeros except in the first row means that \(\alpha_1 = 0\). For the same reason, \(\alpha_2\) is zero. We only have two nonzero rows, and they are linearly independent, so the rank of the matrix is 2.

The span of the rows is called the row space. The row space of \((A)\) and the row echelon form of \((A)\) are the same. In the example, \[
\begin{bmatrix} 1 & 2 & 3 & 4 \\
0 & 0 & -1 & -2
\end{bmatrix}
\]
The span of columns is called the column space. Similarly to row space, the span of columns is called the column space. The span of the rows is called the row space. Similarly to row space, the span of columns is called the column space. The span of the rows is called the row space. Similarly to row space, the span of columns is called the column space.

Theorem \(\PageIndex{1}\)

\[
\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix} = \operatorname{span} \left\{ \begin{bmatrix} 1 \end{bmatrix}, \begin{bmatrix} 2 \end{bmatrix}, \begin{bmatrix} 3 \end{bmatrix} \right\}.
\]

In particular, to find a set of linearly independent columns we need to look at where the pivots were. If you recall above, when solving \((A \vec{x} = \vec{0})\) the key was finding the pivots, any non-pivot columns corresponded to free variables. That means we can solve for the non-pivot columns in terms of the pivot columns. Let’s see an example. First we reduce some random matrix: \[
\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{bmatrix}
\]
We find the next pivot, make it one, and rinse and repeat: \[
\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{bmatrix}
\]
The final matrix is the row echelon form of the matrix. Consider the pivots that we marked. The pivot columns are the first and the third column. All other columns correspond to free variables when solving \((A \vec{x} = \vec{0})\), so all other columns can be solved in terms of the first and the third column. In other words \[
\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \\
1 & 2 & 3 & 4
\end{bmatrix}
\]
The discussion above could be expanded into a proof of the theorem if we wanted. As each nonzero row in the row echelon form contains a pivot, then the rank is the number of pivots, which is the same as the maximal number of linearly independent columns.

The idea also works in reverse. Suppose we have a bunch of column vectors and we just need to find a linearly independent set. For example, suppose we started with the vectors \[ \vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}. \]

These vectors are not linearly independent as we saw above. In particular, the span of \( \langle \vec{v}_1 \rangle \) and \( \langle \vec{v}_3 \rangle \) is the same as the span of all four of the vectors. So \( \langle \vec{v}_2 \rangle \) and \( \langle \vec{v}_4 \rangle \) can both be written as linear combinations of \( \langle \vec{v}_1 \rangle \) and \( \langle \vec{v}_3 \rangle \). A common thing that comes up in practice is that one gets a set of vectors whose span is the set of solutions of some problem. But perhaps we get way too many vectors, we want to simplify. For example above, all vectors in the span of \( \langle \vec{v}_1 \rangle \) and \( \langle \vec{v}_3 \rangle \) can be written \( a \vec{v}_1 + b \vec{v}_3 \) for two numbers \( a \) and \( b \). And one has to admit, that looks much simpler. Moreover, these numbers \( a \) and \( b \) are unique. More on that in the next section.

To find this linearly independent set we simply take our vectors and form the matrix \[ \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix}. \]

We crank up the row-reduction machine, feed this matrix into it, find the pivot columns, and pick those. In this case, \( \langle \vec{v}_1 \rangle \) and \( \langle \vec{v}_3 \rangle \).

**Computing the Inverse**

If the matrix \( A \) is square and there exists a unique solution \( \langle \vec{x} \rangle \) to \( A \vec{x} = \vec{b} \) for any \( \langle \vec{b} \rangle \) (there are no free variables), then \( A \) is invertible. This is equivalent to the \( n \times n \) matrix \( A \) being of rank \( n \).

In particular, if \( A \vec{x} = \vec{b} \) then \( \langle \vec{x} \rangle = A^{-1} \vec{b} \). Now we just need to compute what \( A^{-1} \) is. We can surely do elimination every time we want to find \( A^{-1} \vec{b} \), but that would be ridiculous. The mapping \( \langle A^{-1} \vec{b} \rangle \) is linear and hence given by a matrix, and we have seen that to figure out the matrix we just need to find where \( \langle A^{-1} \vec{b} \rangle \) takes the standard basis vectors \( \langle \vec{e}_1 \rangle, \langle \vec{e}_2 \rangle, \ldots, \langle \vec{e}_n \rangle \).

That is, to find the first column of \( A^{-1} \vec{b} \), we solve \( A \vec{x} = \vec{e}_1 \), because then \( A^{-1} \vec{e}_1 = \vec{x} \). To find the second column of \( A^{-1} \vec{b} \), we solve \( A \vec{x} = \vec{e}_2 \). And so on. It is really just \( n \) eliminations that we need to do. But it gets even easier. If you think about it, the elimination is the same for everything on the
left side of the augmented matrix. Doing \((n)\) eliminations separately we would redo most of the computations. Best is to do all at once.

Therefore, to find the inverse of \((A)\), we write an \((n \times 2n)\) augmented matrix \([ A \sim \mathbf{I} \mathbf{I} ]\), where \((\mathbf{I})\) is the identity matrix, whose columns are precisely the standard basis vectors. We then perform row reduction until we arrive at the reduced row echelon form. If \((A)\) is invertible, then pivots can be found in every column of \((A)\), and so the reduced row echelon form of \(([ A \sim \mathbf{I} \mathbf{I} ])\) looks like \([ [ I \sim A^{-1} ] \mathbf{I} \mathbf{I} ]\). We then just read off the inverse \((A^{-1})\). If you do not find a pivot in every one of the first \((n)\) columns of the augmented matrix, then \((A)\) is not invertible.

This is best seen by example. Suppose we wish to invert the matrix \([ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix} \) . We write the augmented matrix and we start reducing:

\[
\begin{align}
\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \to & \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -4 & -5 & -2 & 1 & 0 \\ 0 & -5 & -9 & -3 & 0 & 1 \end{bmatrix} \to \\
& \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{-11}{4} & \frac{-1}{2} & \frac{-5}{4} & 1 \end{bmatrix} \to \\
& \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & \frac{2}{11} & \frac{5}{11} & \frac{-4}{11} \end{bmatrix} \to \\
& \begin{bmatrix} 1 & 0 & 0 & \frac{-1}{11} & \frac{3}{11} & \frac{2}{11} \\ 0 & 1 & 0 & \frac{3}{11} & \frac{-9}{11} & \frac{5}{11} \end{bmatrix} \\
& \begin{bmatrix} 1 & 0 & 0 & \frac{-1}{11} & \frac{3}{11} & \frac{2}{11} \\ 0 & 1 & 0 & \frac{3}{11} & \frac{-9}{11} & \frac{5}{11} \end{bmatrix} \end{align}
\]

So \([ \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \] = \begin{bmatrix} 1 & 0 & 0 & \frac{-1}{11} & \frac{3}{11} & \frac{2}{11} \\ 0 & 1 & 0 & \frac{3}{11} & \frac{-9}{11} & \frac{5}{11} \end{bmatrix} \] . Not too terrible, no? Perhaps harder than inverting a \((2 \times 2)\) matrix for which we had a simple formula, but not too bad. Really in practice this is done efficiently by a computer.

Footnotes

[1] Although perhaps we have this backwards, quite often we solve a linear system of equations to find out something about matrices, rather than vice versa.