These notes cover the basic facts regarding three fundamental topics in linear algebra:

- Linear independence;
- Bases;
- Dimension.

Throughout these notes, $V$ will denote a vector space. You can imagine that $V$ is one of the vector spaces below:

- $\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in \mathbb{R} \right\}$;
- The vector space $\mathbb{R}^{m \times n}$ consisting of all $m \times n$ matrices with real entries;
- The vector space $\mathbb{P}_n$ consisting of all polynomials of degree at most $n$;
- The vector space $\mathbb{P}$ consisting of all polynomials (of any degree);
- The vector space $C(\mathbb{R})$ consisting of all continuous functions $f : \mathbb{R} \to \mathbb{R}$;
- The vector space $C^\infty(\mathbb{R})$ consisting of all infinitely differentiable functions $f : \mathbb{R} \to \mathbb{R}$.

Note that in each of these examples, we know what it means to add two elements (ordinary vectors, matrices, polynomials, or functions) from a given space, and we know how to multiply these elements by scalars. We will refer to the elements of these vector spaces as vectors, even if they happen to be objects of another sort.

1. Review of spanning sets

We have already discussed the idea of a spanning set for a subspace $S \subset V$. This is a collection of vectors $s_1, \ldots, s_k$ that spans $S$, meaning that every vector in $S$ can be built as a linear combination of the given vectors $s_1, \ldots, s_k$.

**Example 1.** Consider a matrix $A \in \mathbb{R}^{m \times n}$. Then we can write down the parametric form for the solution set of the homogenous system $A\vec{x} = \vec{0}$. The vectors appearing in this parametric form span the nullspace $N(A)$, because every solution to $A\vec{x} = \vec{0}$ is a linear combination of the vectors appearing in the parametric form. (Remember that $N(A)$ is defined to be the set of all solutions to $A\vec{x} = \vec{0}$.)

For instance, if $A = \begin{pmatrix} 1 & 1 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & -3 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$ then the parametric form for the solution set is

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -3 \\ 0 \\ 0 \\ 1 \end{pmatrix} + x_6 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
$$
This equation states exactly what we want: each solution \[
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5 \\
x_6
\end{pmatrix}
\] is a linear combination of the “special” solutions
\[
\begin{pmatrix}
-1 \\
1 \\
0 \\
0 \\
0 \\
0
\end{pmatrix},
\begin{pmatrix}
-3 \\
4 \\
0 \\
0 \\
0 \\
0
\end{pmatrix},
\begin{pmatrix}
0 \\
0 \\
1 \\
0 \\
0 \\
1
\end{pmatrix}.
\]

(Recall that these “special” solutions are obtained by setting one free variable equal to 1, and the others equal to zero.)

2. Linear Independence

Idea. Say we have a spanning set for a subspace \( S \subset V \). This spanning set lets us describe all vectors in \( S \) (by building combinations) but is this the most efficient way to describe \( S \)? We would like to make sure that there are no redundancies among the vectors in our set.

Example 2. Consider the vectors
\[
\begin{pmatrix}
-1 \\
1 \\
0 \\
0 \\
0 \\
0
\end{pmatrix},
\begin{pmatrix}
-3 \\
4 \\
0 \\
0 \\
0 \\
0
\end{pmatrix},
\begin{pmatrix}
0 \\
0 \\
1 \\
0 \\
0 \\
1
\end{pmatrix}.
\]

These vectors form a spanning set for the nullspace \( N(A) \) (where \( A \) is the matrix from the Example 1), but they are redundant: the last vector is simply the sum of the previous three, and hence we don’t really need to use it when building vectors in \( N(A) \).

Linear independence is a condition that rules out such redundancy.

Definition 2.1. We say that a collection of vectors \( \vec{v}_1, \ldots, \vec{v}_k \) in \( V \) is linearly independent (or simply independent) if the only scalars \( \alpha_1, \ldots, \alpha_k \) that satisfy the equation
\[
\alpha_1 \vec{v}_1 + \alpha_1 \vec{v}_2 + \cdots + \alpha_k \vec{v}_k = \vec{0}
\]
are \( \alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_k = 0 \).

Another way to say this: the only way to build the zero vector \( \vec{0} \) out of the vectors \( \vec{v}_i \) is the trivial way.

Example 3. The vectors \( \vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \) and \( \vec{v}_2 = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} \) and not linearly independent, because \( \vec{v}_2 = 2\vec{v}_1 \) and hence \( 2\vec{v}_1 - \vec{v}_2 = \vec{0} \) gives a non-trivial way to build the zero vector out of \( \vec{v}_1 \) and \( \vec{v}_2 \).
Example 4. Say \( V = \mathbb{R}^m \). Then vectors \( \vec{v}_1, \ldots, \vec{v}_n \) in \( \mathbb{R}^m \) are linearly independent if and only if the homogeneous equation

\[
[\vec{v}_1 \vec{v}_2 \ldots \vec{v}_m] \vec{x} = \vec{0}
\]

has only the trivial solution \( \vec{x} = \vec{0} \). Here \( [\vec{v}_1 \vec{v}_2 \ldots \vec{v}_m] \) is the \( m \times n \) matrix with \( i \)th column \( \vec{v}_i \).

To see why this is true, remember that multiplying a matrix by a vector results in a linear combination of the columns of the matrix.

You can now check whether vectors in \( \mathbb{R}^n \) are linearly independent using methods from earlier in the course, because we’ve turned the problem into a problem about linear systems.

Exercise 1. Which of the following lists of vectors are linearly independent? (Note: you should be able to answer some of these without actually doing any row reduction.)

\[
\begin{align*}
\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 9 \end{pmatrix}, \begin{pmatrix} 1 \\ 7 \end{pmatrix} \right\}, & \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 9 \end{pmatrix}, \begin{pmatrix} 1 \\ 7 \end{pmatrix} \right\} \\
\left\{ \begin{pmatrix} 2 \\ 8 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \end{pmatrix} \right\}, & \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.
\end{align*}
\]

What does the condition in Definition 2.1 have to do with the idea of redundancy? With a little bit of work you can check that this condition is the same as requiring the no vector from the set can be built as a combination of the other vectors in the set. In other words, a set is linearly independent if and only if there are no “redundancies.”

We will say that a set of vectors is linearly dependent if it is not linearly independent. In other words a set of vectors is dependent if it contains a redundancy: one vector in the list can be built from the others.

Important: Here redundancy does not mean that a vector is listed twice. For example, the collection of vectors in Example 2 is linearly dependent, but no vector appears twice in the list.

3. Bases

Definition 3.1. A basis for a vector space \( V \) is a linearly independent collection of vectors that spans \( V \).

What does this mean? If \( B = \{ \vec{v}_1, \ldots, \vec{v}_n \} \) is a basis for \( V \), then every vector \( v \in V \) can be built from the vectors \( \vec{v}_i \) in exactly one way:

- The fact that \( B \) is a spanning set means that every vector in \( V \) can be built from the vectors \( \vec{v}_i \);
- If there were two different ways to build a vector \( \vec{v} \in V \) out of the vectors \( \vec{v}_i \), then there would be a non-trivial way to build \( \vec{0} \) out of the \( \vec{v}_i \): simply subtract the two equations that build \( \vec{v} \).
Example 5. The vectors $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, and $\vec{v}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ span $\mathbb{R}^3$ (check!) but do not form a basis for $\mathbb{R}^3$. These vectors can be used to build the vector $\begin{pmatrix} 3 \\ 7 \end{pmatrix}$ in two different ways:

$$3\vec{v}_1 + \vec{v}_3 = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$$

and

$$\vec{v}_1 + \vec{v}_2 + 2\vec{v}_3 = \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$ 

By subtracting these equations, we see that

$$2\vec{v}_1 - \vec{v}_2 - \vec{v}_3 = \vec{0}.$$ 

So $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is not linearly independent.

Example 6.

- The standard vectors $\vec{e}_1, \ldots, \vec{e}_n$ for a basis for $\mathbb{R}^n$.
- The vectors appearing in the parametric form for the solution set to a homogeneous system $A\vec{x} = \vec{0}$ form a basis for $N(A)$.
- If $R$ is the reduced echelon form of $A$, then the non-zero rows of $R$ form a basis for the row space $\text{Row}(A)$.
- The pivot columns of a matrix $A$ form a basis for the column space $\text{Col}(A)$.
- The columns of an invertible $n \times n$ matrix $A$ form a basis for $\mathbb{R}^n$.

Each of these facts follows quickly from things we already know. For example, in each of the above cases, we already know that the vectors in question span the space in question (think about this!). So it remains to figure out why these sets are linearly independent. The first four are very similar. For example, the vectors in the parametric form from Example 1 are linearly independent, because if you build a combination of the vectors

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -3 \\ 4 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \\ 2 \end{pmatrix}$$

using the coefficients $x_2, x_4,$ and $x_6$ (where the names of the coefficients are chosen to correspond to the free variables in the original system) you get

$$x_2\vec{v}_1 + x_4\vec{v}_2 + x_6\vec{v}_3 = \begin{pmatrix} -x_2 - 3x_4 \\ x_2 \\ 3x_6 \\ x_4 \\ 2x_6 \end{pmatrix},$$

and by looking at the second, fourth, and sixth entries of the vector on the right, we see that the only way to build the zero vector out of $\vec{v}_1, \vec{v}_2,$ and $\vec{v}_3$ is to set $x_2 = x_4 = x_6 = 0$. 

Said another way, the only way to build the trivial solution to a homogeneous system is to set all the free variables to zero!

Exercise 2. Explain why the columns of an invertible $n \times n$ matrix must form a basis for $\mathbb{R}^n$. (Think about this in terms of solutions to linear systems.)

4. Dimension

Definition 4.1. The dimension of a vector space $V$, written $\dim(V)$, is defined to be the number of elements in a basis for $V$. Note that this number could be infinity.

We now state one of the most important facts in linear algebra. (This theorem tells us that the previous definition makes sense.)

Theorem 4.2. If $V$ is a vector space and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for $V$, then every basis for $V$ has exactly $n$ elements.

A more detailed version of this statement is given at the end of the notes.

Example 7.

• Since the standard vectors $\vec{e}_i$ form a basis for $\mathbb{R}^n$, we have $\dim(\mathbb{R}^n) = n$ (as you’d expect).
• The polynomials $1, x, x^2, \ldots, x^n$ form a basis for $\mathbb{P}_n$, so $\dim(\mathbb{P}_n) = n + 1$.
• The polynomials $1, x, x^2, \ldots, x^n, x^{n+1}, \ldots$ form a basis for $\mathbb{P}$, so $\dim(\mathbb{P}) = \infty$.
• The $m \times n$ matrices $E(ij)$ with entries $E(ij)_{ij} = 1$ and all other entries equal to zero form a basis for $\mathbb{R}^{m \times n}$, so $\dim \mathbb{R}^{m \times n} = mn$.

There is no natural way to describe bases for the vector spaces $C(\mathbb{R})$ and $C^\infty(\mathbb{R})$. However, these vector spaces contain the vector space $\mathbb{P}$ (since polynomials are infinitely differentiable) so these spaces are also infinite dimensional.

Example 8: dimensions of the fundamental subspaces. If $A$ is an $m \times n$ matrix, then

• $\dim N(A)$ is the number of free variables in the system $A\vec{x} = 0$ (i.e. the number of non-basic columns of $A$);
• $\dim \text{Col}(A)$ is the number of pivot variables (i.e. the number of basic columns);
• $\dim \text{Row}(A)$ is the number of pivot rows in the reduced echelon form of $A$, so $\dim \text{Row}(A) = \dim \text{Col}(A)$;
• $\dim N(A^T) = m - \text{rank}(A)$. We will explain this below.

Theorem 4.3 (Rank–Nullity Theorem). For any $m \times n$ matrix $A$,

$$\dim N(A) + \dim \text{Col}(A) = n.$$

In terms of Example 8, this theorem simply says that the number of free variables in the system $A\vec{x} = 0$, plus the number of pivot variables, equals the total number of variables.

Example 9. Consider an $m \times n$ matrix $A$. We’d like to understand the left-hand nullspace $N(A^T)$. Since $A^T$ is an $n \times m$ matrix, the Rank–Nullity Theorem says that

$$\dim N(A^T) + \dim \text{Row}(A) = m.$$
Since we know that \( \dim \text{Row}(A) = \dim \text{Col}(A) = \rank(A) \) (all three of these numbers are the number of pivots in the \( R \), reduced echelon form of \( A \)), this gives
\[
\dim N(A^T) = m - \rank(A),
\]
as claimed in Example 8. Notice that this is the number of rows in \( R \) minus the number of non-zero rows in \( R \), so we have
\[
\dim N(A^T) = \text{number of rows of zeros in } R.
\]

5. Spanning and independence in \( n \)-dimensional space

We conclude these notes with a more detailed version of Theorem 4.2.

**Theorem 5.1.** If \( V \) is an \( n \)-dimensional vector space, then every set of more than \( n \) vectors in \( V \) is linearly dependent, and no set of fewer than \( n \) vectors spans \( V \).

In addition, every independent set of \( n \) vectors in \( V \) actually spans \( V \) (hence forms a basis), and every spanning set of \( n \) vectors in \( V \) is actually independent (hence forms a basis). This means that if \( W \subset V \) is a subspace and \( W \) is \( n \)-dimensional, then in fact \( W = V \).

Although we won’t go through a detailed proof, here are the basic ideas. If \( V \) is \( n \)-dimensional, we can choose a basis \( \vec{v}_1, \ldots, \vec{v}_n \) for \( V \).

If \( \{\vec{w}_1, \ldots, \vec{w}_m\} \) are vectors in \( V \), then it’s possible to express all of the \( \vec{w}_i \) in terms of the vectors \( \vec{v}_i \). If you put all the coefficients that you’ve used into a matrix \( A \), (where the coefficients for \( \vec{w}_1 \) go in the first column, etc.) then you end up with an \( n \times m \) matrix. For example, say \( m = 4 \) and \( n = 3 \). If \( \vec{w}_1 = 2\vec{v}_1 + 3\vec{v}_2, \vec{w}_2 = \vec{v}_2 + \vec{v}_3, \vec{w}_3 = \vec{v}_1 - \vec{v}_2 - \vec{v}_3 \) and \( \vec{w}_4 = \vec{v}_2 - 2\vec{v}_3 \), the matrix is
\[
A = \begin{pmatrix}
2 & 0 & 1 & 3 \\
3 & 1 & -1 & 1 \\
0 & 1 & -1 & -2
\end{pmatrix}.
\]

If \( m \) is more than \( n \) (as in the example), the columns of \( A \) are dependent (the homogeneous system \( A\vec{x} = \vec{0} \) has a free variable). This means that there is a non-trivial linear relation among the columns of \( A \), and in fact the same relation holds among the vectors \( \vec{w}_i \), and this makes the vectors \( \vec{w}_i \) dependent. So sets in \( V \) containing more than \( n \) vectors are dependent. To see how this works in the example, note that if we write \( A = [\vec{a}_1 \vec{a}_2 \vec{a}_3 \vec{a}_4] \), then you can check that \( \vec{a}_4 = \vec{a}_1 - \vec{a}_2 + \vec{a}_3 \).

You can also check that \( \vec{w}_4 = \vec{w}_1 - \vec{w}_2 + \vec{w}_3 \) (this computation is almost exactly the same).

If \( \vec{w}_1, \ldots, \vec{w}_m \) span \( V \), then it’s possible to express all of the \( \vec{v}_i \) in terms of the \( m \) vectors \( \vec{w}_1, \ldots, \vec{w}_m \). If you put all the coefficients that you’ve used into a matrix \( A \), (where the coefficients for \( \vec{v}_1 \) go in the first column, etc.). This matrix has \( n \) columns and \( m \) rows. If \( n \) were actually more than \( m \), the columns would be linearly dependent, giving a non-trivial linear relation among the columns of \( A \). The same linear relation would hold among the vectors \( \vec{v}_i \) (just as in the previous paragraph). But this would make the vectors \( \vec{v}_i \) dependent, which is impossible since they form a basis.

Finally, consider a set \( \vec{w}_1, \ldots, \vec{w}_n \) of vectors in \( V \). We can again express these vectors in terms of the basis \( \vec{v}_1, \ldots, \vec{v}_n \), and we obtain an \( n \times n \) matrix \( A \). The ideas earlier in these notes show that the columns of an \( n \times n \) matrix are linearly independent if and only if they span \( \mathbb{R}^n \) (why?). This can now be translated into the statement that the vectors \( \vec{w}_i \) are independent if and only if they span \( V \), as desired.

For the very last statement in the Theorem, say \( W \subset V \) is an \( n \)-dimensional subspace of \( V \). Then \( W \) has a basis containing \( n \) vectors, and this linearly independent set of \( n \) vectors must actually span \( V \). So every vector \( v \) in \( V \) can be built from vectors in \( W \). Since \( W \) is a subspace this means every \( v \in V \) is also an element of \( W \), so we must have \( V = W \).