4.3 Linearly Independent Sets; Bases

Definition An indexed set $\{\bar{v}_1, \ldots, \bar{v}_p\}$ in a vector space V is said to be **linearly independent** if

$$c_1\bar{v}_1 + \cdots + c_p\bar{v}_p = \bar{0}$$

has only the trivial solution $0 = c_1 = \cdots = c_p$. The set $\{\bar{v}_1, \ldots, \bar{v}_p\}$ is **linearly dependent** if there is a nontrivial solution to this equation.

The following theorem should look familiar, but in this case we are talking about vector spaces in general, rather than just $\mathbb{R}^n$.

Theorem An indexed set $\{\bar{v}_1, \ldots, \bar{v}_p\}$ of two or more vectors in a vector space V, with $\bar{v}_1 \neq \bar{0}$, is linearly dependent if and only if some $\bar{v}_j$ (for some $j > 1$) is a linear combination of the preceding vectors $\bar{v}_1, \ldots, \bar{v}_{j-1}$.

Definition Let H be a subspace of a vector space V. An indexed set of vectors $\mathcal{B} = \{\bar{b}_1, \ldots, \bar{b}_p\}$ in V is a **basis** for H if
(i) $\mathcal{B}$ is a linearly independent set, and
(ii) $H = \text{span}\mathcal{B} = \text{span}\{\bar{b}_1, \ldots, \bar{b}_p\}$.

In words, a basis for a vector (sub)space is a linearly independent set of vectors that spans the (sub)space.

Example If A is an invertible $n \times n$ matrix, then the columns of A form a basis of $\mathbb{R}^n$. By the Invertible Matrix Theorem, the columns are linearly independent and they span $\mathbb{R}^n$, which is exactly the definition of a basis of $\mathbb{R}^n$.

Example As a special case of the previous example, consider the $n \times n$ identity matrix I_n. This matrix is invertible, so its columns are a basis of $\mathbb{R}^n$. In fact, its columns are $\bar{e}_1, \ldots, \bar{e}_n$, which we call the **standard basis** of $\mathbb{R}^n$. In the case of $\mathbb{R}^2$, the standard basis is $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$.

Example We determine if

$$\left\{ \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} \right\}$$

is a basis for $\mathbb{R}^3$. We row reduce the matrix whose columns are formed by the vectors:

$$\begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ -6 & 7 & 5 \end{bmatrix} \sim \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

So we see there are three pivots, and the matrix is therefore invertible. It follows that the vectors are a basis of $\mathbb{R}^3$.

Example The **standard basis** of $\mathbb{P}_n$ (polynomials of degree at most n) is

$$\{1, x, x^2, \ldots, x^n\}.$$

To see why, note first of all that this set spans $\mathbb{P}_n$, since each element of $\mathbb{P}_n$ can (by definition) be written

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x^1 + a_0 1.$$

It remains to show why the set is linearly independent. For this, note that if

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x^1 + a_0 1 = 0,$$

then $p(x)$ must be the zero polynomial. Why? If not, then by the Fundamental Theorem of Algebra, $p(x)$ would have *at most* n roots, which means the equation $p(x) = 0$ would not be satisfied for all x. It follows that $0 = a_0 = a_1 = \cdots = a_n$; i.e. the only solution is the trivial solution.

Theorem (The Spanning Set Theorem) Let $S = \{\bar{v}_1, \dots, \bar{v}_p\}$ be a set in V, and let $H = \text{Span}\{\bar{v}_1, \dots, \bar{v}_p\}$.
(a) If one of the vectors in S – say, $\bar{v}_k$ – is a linear combination of the others, then the set formed by removing $\bar{v}_k$ from S still spans H.
(b) If $H \neq \{\bar{0}\}$, then some subset of S is a basis for H.

The intuition for why this theorem is true is more important than its proof. For (a), the basic idea is that any vector that is a linear combination of the others is extraneous information. For (b), we apply (a) repeatedly until there is no more extraneous information, at which point we are left with a linearly independent set that still spans H (hence is a basis).

An important thing to understand about bases is that they satisfy a *minimality* property and a *maximality* property. How so? The basis of a vector space is *minimal as a spanning set.* That is, if you remove any vector, the set no longer spans the space. It is *maximal as a linearly independent set*; i.e. if you add any vector, the set is no longer linearly independent.

Bases for Null Spaces and Column Spaces If A is an $m \times n$ matrix, we saw that Nul$A$ is the set of vectors $\bar{x}$ satisfying

$$A\bar{x} = \bar{0},$$

and that these vectors are a subspace of $\mathbb{R}^n$. When writing the solution set in parametric vector form, the vectors are linearly independent, and since they also span Nul$A$, they provide a ***basis*** for Nul$A$ as well. For finding a basis of Col$A$, we have the following theorem.

Theorem The pivot columns of a matrix A form a basis for Col$A$.

Example We find a basis for Col$B$, where

$$B = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The pivot columns are 1, 3, and 5, so a basis for Col$B$ is

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Example Let A be the matrix

$$A = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}.$$

It can be shown that A is row equivalent to B from the preceding example. In other words, the solution sets of $A\bar{x} = \bar{0}$ and $B\bar{x} = \bar{0}$ are the same, from which it follows that the same linear dependencies that hold for Col$B$ also hold for Col$A$. Hence columns 1, 3, and 5 form a basis for Col$A$. That is, a basis for Col$A$ is

$$\left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 2 \\ 8 \end{bmatrix} \right\}.$$

4.4 Coordinate Systems

Theorem (The Unique Representation Theorem) Let $\mathcal{B} = \{\bar{b}_1, \ldots, \bar{b}_n\}$ be a basis for a vector space V. Then for each $\bar{x} \in V$, there is a unique set of scalars $c_1, \ldots, c_n$ such that

$$\bar{x} = c_1\bar{b}_1 + \cdots + c_n\bar{b}_n.$$

We say that the weights $c_1, \ldots, c_n$ are the **$\mathcal{B}$-coordinates of $\bar{x}$.**

Definition If $c_1, \ldots, c_n$ are the $\mathcal{B}$-coordinates of $\bar{x}$, then the vector

$$[\bar{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

is the **$\mathcal{B}$-coordinate vector of** $\bar{x}$. We call the map

$$\bar{x} \mapsto [\bar{x}]_{\mathcal{B}}$$

the **coordinate mapping of** $\mathcal{B}$.

So what is this all saying? Basically, for any basis of V, we can represent each vector in V in terms of the basis.

Example Consider the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ of $\mathbb{R}^2$. Let $\bar{x}$ be a vector in $\mathbb{R}^2$ with $[\bar{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. It follows that

$$\bar{x} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Note This is itself a representation of the vector $\bar{x}$ with respect to the *standard basis* $\mathcal{E} = \{\bar{e}_1, \bar{e}_2\}$. It follows that that $[\bar{x}]_{\mathcal{E}} = \bar{x}$ for all $\bar{x} \in V$.

Geometrically, you can think of different bases for a vector spaces as corresponding to different lattices on the space. We will draw some examples in class.

Coordinates in $\mathbb{R}^n$ If we fix a basis $\mathcal{B}$ of $\mathbb{R}^n$, then it is a straightforward exercise to find $[\bar{x}]_{\mathcal{B}}$ for all $\bar{x}$.

Example Let $\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^2$ (confirm this!). To find the coordinate vector

$$\begin{bmatrix} 4 \\ 5 \end{bmatrix}_{\mathcal{B}},$$

we need weights c_1 and c_2 such that

$$c_1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

In other words, we are solving the matrix equation

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

At this point, we have several methods for solving such a matrix equation. You can pick your favorite and work out the answer on your own. You should find that $c_1 = 3$ and $c_2 = 2$ so that

$$\begin{bmatrix} 4 \\ 5 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

The matrix

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

in the preceding example is what we call a **change-of-coordinates** matrix. Specifically, it takes a vectors written in terms of $\mathcal{B}$ and turns it into a vector written in terms of $\mathcal{E}$. More generally, for any basis $\mathcal{B} = \{\bar{b}_1, \ldots, \bar{b}_n\}$ of $\mathbb{R}^n$, the matrix

$$P_{\mathcal{B}} = \begin{bmatrix} \bar{b}_1 & \bar{b}_2 & \ldots & \bar{b}_n \end{bmatrix}$$

is the matrix that turns a vector written in terms $\mathcal{B}$ into a vector written in terms of the standard basis $\mathcal{E}$ via the map

$$[\bar{x}]_{\mathcal{B}} \mapsto P_{\mathcal{B}}[\bar{x}]_{\mathcal{B}}.$$

Note that since $\mathcal{B}$ is a basis, the matrix $P_{\mathcal{B}}$ is invertible. It follows that for any vector $\bar{x}$ written in terms of the standard basis,

$$P_{\mathcal{B}}^{-1}\bar{x} = [\bar{x}]_{\mathcal{B}}.$$

Theorem Let $\mathcal{B} = \{\bar{b}_1, \ldots, \bar{b}_n\}$ be a basis for a vector space V. Then the coordinate mapping $\bar{x} \mapsto [\bar{x}]_{\mathcal{B}}$ is a one-to-one linear transformation from V onto $\mathbb{R}^n$.

In general, we call a one-to-one and onto linear map of vectors spaces an *isomorphism*, because the two spaces have algebraically the "same structure."

Example Working in $\mathbb{P}_n$ is the same as working in $\mathbb{R}^{n+1}$. To see why, note that an element $p(x) \in \mathbb{P}_n$ written in terms of the standard basis $\mathcal{E} = \{1, x, \ldots, x^n\}$ has the form

$$p(x) = a_0 + a_1 x^1 + \cdots + a_{n-1}x^{n-1} + a_n x^n.$$

It follows that

$$[p(x)]_{\mathcal{E}} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix} \in \mathbb{R}^{n+1}.$$

In other words, the map

$$p(x) \mapsto [p(x)]_{\mathcal{E}}$$

is an isomorphism of $\mathbb{P}_n$ onto $\mathbb{R}^{n+1}$.