Bases

From your study of calculus, you are familiar with the fact that you can describe *any* point in the xy plane (i.e., in $\mathbb{R}^2$) in terms of the vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix};$$

for example, we can describe the vector

$$\begin{pmatrix} 2 \\ -5 \end{pmatrix}$$

as

$$2e_1 - 5e_2.$$

Of course, we must have *both* e_1 and e_2 in order to do so–if we "threw away" one of the two vectors, we would no longer be able to describe many points in the xy plane.

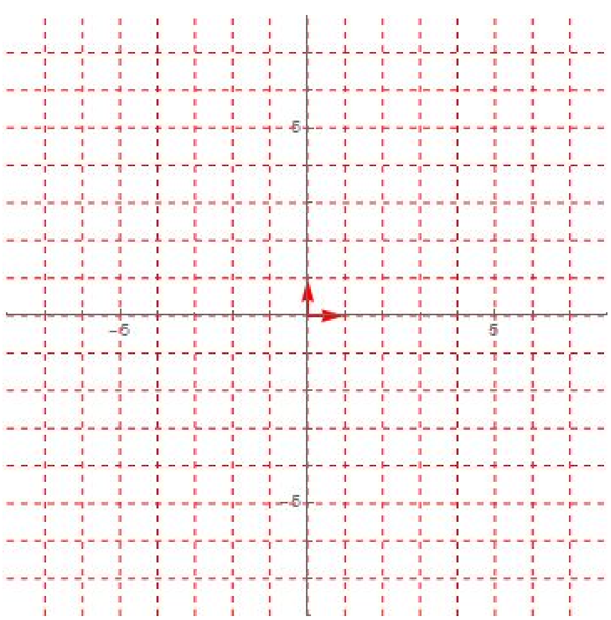
There is something essential about the list $B = (e_1, e_2)$ in $\mathbb{R}^2$: it includes enough vectors so that we can describe all of $\mathbb{R}^2$ using linear combinations of its vectors (i.e., B spans $\mathbb{R}^2$); but it does not include any unnecessary vectors–B is a linearly independent set.

However, we have also seen other ways to describe points in the plane, say by using the polar coordinates that you studied in calculus. We could even use the vectors

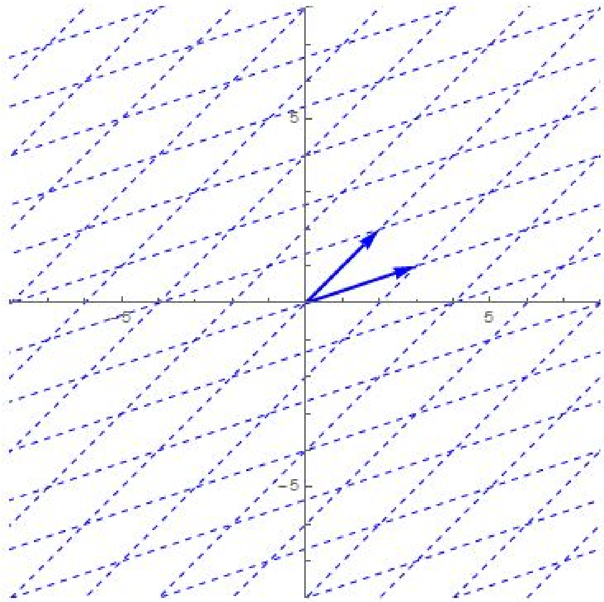
$$v_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

that we discussed in the last section–indeed, we saw that these two vectors can be used to describe any point in $\mathbb{R}^2$.

In a sense, we now have multiple ways to think about the geometry or "shape" of the plane: we can describe points as usual, thinking of them in terms of the vectors e_1 and e_2:



Alternatively, we can think about points in the plane in terms of the vectors v_1 and v_2:



Both systems of representing points in the plane are perfectly acceptable: both systems give us a way to describe the location of points, and both systems give us a way to measure the distance between points.

In this section, we will make these ideas more precise by introducing *bases* and *coordinate systems*. Our goal is to make the preceding discussion more general: we want to develop ways to describe the "geometry" of other vector spaces: to talk about the locations of vectors in the space, and to describe the distances between vectors.

Bases

We mentioned above that the list $B = (e_1,\ e_2)$ is important in $\mathbb{R}^2$ because it includes enough vectors so that we can describe all of $\mathbb{R}^2$ (B spans $\mathbb{R}^2$), but it does not include any unnecessary vectors (B is linearly independent).

We will see soon that a list S of vectors from a vector space V that is *both* linearly independent *and* spanning has special properties; we give such a subset a name:

Definition 2.27. A list S of vectors from a vector space V is called a *basis* for V if:

1. S is independent, and
2. $\text{span}\,(S) = V$.

The following theorem indicates the reason that we go to the trouble to define a basis in the first place:

Theorem 2.29. A list $S = (v_1,\ v_2,\ \ldots,\ v_n)$ is a basis for vector space V if and only if every vector v in V can be written uniquely as a linear combination of the basis vectors in S.

Remark. The theorem is particularly helpful: it says that (1) every $v \in V$ can be written as a linear combination of vectors in S, and (2), if v can be written as

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n,$$

then there is *no other* linear combination of the basis vectors that will yield v.

You can think of a basis for a vector space as a set that describes everything essential about the space, without introducing ambiguity: since a basis spans the space, it includes enough vectors to describe the entire space; but since the basis is also linearly independent, it does not include any unnecessary vectors.

Examples of Bases

Example 1: $\mathbb{R}^n$

We have seen that the vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

span $\mathbb{R}^2$ and are linearly independent; thus they form a basis for $\mathbb{R}^2$. Incidentally, we have also seen that the vectors

$$v_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

form a basis for $\mathbb{R}^2$.

In general, the vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\ \text{and } e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ n \end{pmatrix}$$

in $\mathbb{R}^n$ span $\mathbb{R}^n$ and are linearly independent; we call them the *standard basis vectors*. However, you should be aware that there are many different ways to choose a basis for $\mathbb{R}^n$.

Example 2: $\mathcal{M}_{mn}$

The vector space $\mathcal{M}_{mn}$ of all $m \times n$ matrices with entries in $\mathbb{F}$ has a basis consisting of $m \cdot n$ vectors: to build these vectors, let e_{ij} be the $m \times n$ matrix whose i, j entry is 1, and all of whose other entries are 0s. This list of vectors spans $\mathcal{M}_{mn}$ and is linearly independent, thus forms a basis form $\mathcal{M}_{mn}$.

As a specific example, consider the vector space $\mathcal{M}_{32}$ of all 3×2 real matrices. This vector space has a basis consisting of the 6 vectors

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad e_{31} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad e_{32} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

We will not show that these vectors span $\mathcal{M}_{32}$, or that they are linearly independent, but both facts are quite easy to see.

Example 3: $\mathcal{P}_n(\mathbb{F})$

Recall that $\mathcal{P}_n(\mathbb{F})$ is the vector space of all polynomials of degree at most n with coefficients in $\mathbb{F}$. It is clear that the list

$$(1, x, x^2, \ldots, x^n)$$

spans $\mathcal{P}_n(\mathbb{F})$; it is not particularly difficult to show that this set is also linearly independent. Thus it forms a basis for $\mathcal{P}_n(\mathbb{F})$.

Example 4: $\mathcal{U}_2(\mathbb{R})$

The vector space $\mathcal{U}_2(\mathbb{R})$ of all real upper triangular 2×2 matrices is spanned by the three matrices

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ and } e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix};$$

these matrices are also linearly independent, and thus the list (e_{11}, e_{12}, e_{22}) is a basis for $\mathcal{U}_2(\mathbb{R})$.

Example 5: $\mathfrak{sl}(2, \mathbb{R})$

The list

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right)$$

spans $\mathfrak{sl}(2, \mathbb{R})$ and is linearly indpendent, thus forms a basis for $\mathfrak{sl}(2, \mathbb{R})$.

Proof of Theorem 2.29. $\Longrightarrow$ On one hand, if S is a basis, and in particular a spanning list for V, then every vector in V is a linear combination of vectors in S. It remains to show that such a linear combination is unique.

Let $v \in V$, and let

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

and

$$v = \beta_1 v_1 + \ldots + \beta_n v_n$$

be any two decompositions of v as a linear combination of vectors in S. Then

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_n v_n$$

implies that

$$(\alpha_1 - \beta_1) v_1 + \ldots + (\alpha_n - \beta_n) v_n = \mathbf{0}.$$

Now S is a basis, and in particular is an independent list. By definition,

$$\alpha_i - \beta_i = 0 \ \forall \ i,$$

so that $\alpha_i = \beta_i$; the decomposition of v is thus unique.

$\Longleftarrow$ On the other hand, suppose that every vector in V is a unique linear combination of vectors in S. Then clearly $\operatorname{span}(S) = V$, and we only need to show independence.

If $\alpha_1, \ldots, \alpha_n$ are constants so that

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = \mathbf{0},$$

then by the assumption that all such decompositions are unique, we must have $\alpha_i = 0$ for all i. The list is thus independent.

Building Bases

In a sense, finite-dimensional vector spaces are particularly nice mathematical structures: if have a basis for the space, then we understand everything we need to now about the space: we can build any vector we like with no ambiguity.

On a related note, finite dimensional vector spaces are nice because it is relatively easy to find a basis for the space, as indicated by the following theorems:

Theorems 2.31/2.33. Let V be a finite dimensional vector space, and let S be a list of vectors from V. Then:

1. If S is a spanning list, then it contains a basis for V.
2. If S is a linearly independent list, then it may be extended to a (finite) basis for V.

Before we look at a proof of the theorem, let us consider one of its many nice consequences: if V is finite-dimensional, then it contains a finite spanning list. By the theorem, we know that this finite spanning list contains a basis. Thus we resolve a difficulty that may have worried you:

Theorem 2.32. Every finite dimensional vector space has a basis.

Proof of Theorems 2.31/2.33. Let V be finite dimensional.

1. Let
$$S = (v_1,\ v_2,\ \ldots,\ v_n)$$
be a finite spanning list; we may assume that S is dependent, because otherwise it is already a basis. In addition, remove any $\mathbf{0}$ vectors, since these do not contribute to $\operatorname{span}(S)$. Since $\operatorname{span}(S) = V$, we may use the Linear Dependence Lemma to delete a vector that is a linear combination of the others, without changing the span of S. We may assume that v_n is a linear combination of the remaining vectors in S (again, reorder the list if not). The list
$$S_1 = (v_1,\ v_2,\ \ldots,\ v_{n-1})$$
spans V, and if it is not independent, we may repeat the process of reordering the list so that the last vector is a linear combination of the previous ones, and deleting the last vector. This method of deleting vectors using the Linear Dependence Lemma ensures that every new list S_i still spans V, and if S_i is also independent, then S_i is a basis. If the process reaches the list
$$S_{n-1} = (v_1),$$
then S_{n-1} is clearly an independent list and forms a basis for V.

2. Let
$$S = (v_1,\ v_2,\ \ldots,\ v_n)$$
be an independent list. By part 1 of the theorem, there is a finite basis
$$S' = (u_1,\ u_2,\ \ldots,\ u_m)$$
for V. If every vector in S' may be written as a linear combination of vectors in S, then *every* vector in V is a linear combination of vectors in S, and S is a basis as well.

   Otherwise, suppose that u_1 is not a linear combination of vectors in S (again, reorder if necessary). Then consider the list
$$S_1 = (u_1,\ v_1,\ v_2,\ \ldots,\ v_n).$$

If there are scalars so that

$$\beta_1 u_1 + \alpha_1 v_1 + \ldots + \alpha_n v_n = \mathbf{0},$$

then we must have $\beta_1 = 0$ since u_1 is not a linear combination of the vectors in S. This implies that $\alpha_i = 0$ for all i since S is independent. Thus S_1 is independent as well. If S_1 also spans V, then it is a basis; otherwise, there is a vector in $S' = (u_1,\ u_2,\ \ldots,\ u_m)$ that is not a linear combination of vectors in S_1, say u_2, and we add it to S_1 to create the independent list

$$S_2 = (u_1,\ u_2,\ v_1,\ v_2,\ \ldots,\ v_n).$$

Using the same logic, we continue the process: if S_2 is not a basis, add a vector from S' that is not in $\operatorname{span}(S_2)$. Of course, S' is finite, so the process must end, and S' spans V, so some S_i must span V as well.

Example. Extend the independent list

$$B = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix},\ \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix} \right)$$

to a basis for $\mathbb{R}^4$.

The list is independent, but clearly not spanning; following the algorithm in the proof of the theorem, we attempt to add vectors from the standard basis

$$(e_1,\ e_2,\ e_3,\ e_4)$$

(where e_i is the vector in $\mathbb{R}^4$ with a 1 in the ith entry and 0s elsewhere) to our independent list.

We can check to see if, say, e_1 is a linear combination of vectors in B by using matrix arithmetic. We wish to determine if there are any solutions to the system

$$a \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The system has augmented matrix

$$\begin{pmatrix} 1 & 3 & | & 1 \\ 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \\ -1 & 2 & | & 0 \end{pmatrix}.$$

We apply elementary row operations to the system to reduce it and find solutions:

$$\begin{pmatrix} 1 & 3 & | & 1 \\ 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \\ -1 & 2 & | & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & | & 1 \\ 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \\ 0 & 5 & | & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 3 & | & 1 \\ 0 & -3 & | & -1 \\ 0 & 1 & | & 0 \\ 0 & 5 & | & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 3 & | & 1 \\ 0 & 1 & | & 0 \\ 0 & -3 & | & -1 \\ 0 & 5 & | & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 3 & | & 1 \\ 0 & 1 & | & 0 \\ 0 & 0 & | & -1 \\ 0 & 5 & | & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 3 & | & 1 \\ 0 & 1 & | & 0 \\ 0 & 0 & | & -1 \\ 0 & 0 & | & 1 \end{pmatrix}.$$

Clearly the system is inconsistent, so e_1 is not a linear combination of the original vectors in B. Thus

$$B_1 = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$$

is an independent list.

Now B_1 may not be a basis; we must check to see if any of the other standard basis vectors are independent from B_1 (and if they are, add them to the list). We check e_2 next, in a slightly different way (to exercise our thinking skills): if e_2 is independent, then the system

$$a \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix} + c \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

has only the trivial solution

$$a = b = c = d = 0.$$

The coefficent matrix in the matrix equation

$$A\mathbf{x} = \mathbf{0}$$

is

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

Before we proceed, let us recall a theorem from Unit 1, Section 10:

Theorem. Let A be an $n \times n$ matrix. Then the following are equivalent:

- A is invertible.
- $Ax = \mathbf{0}$ has only the trivial solution.
- The reduced row echelon form of A is I_n.
- $Ax = b$ is consistent for every $n \times 1$ matrix b.
- $Ax = b$ has exactly one solution for every $n \times 1$ matrix b.
- $\det A \neq 0$.

We can use this powerful theorem to help us determine whether or not the original system has nontrivial solutions: in particular, there are nontrivial solutions *if and only if* $\det A \neq 0$. Let's calculate the determinant; we proceed by expanding along column 4, followed by column 3:

$$\begin{aligned} \det \begin{pmatrix} 1 & 3 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \end{pmatrix} &= 1 \cdot \det \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 2 & 0 \end{pmatrix} \\ &= 1 \cdot \det \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \\ &= 1. \end{aligned}$$

A has nonzero determinant, implying that the system

$$A\mathbf{x} = \mathbf{0}$$

has *only* the trivial solution, thus that the list

$$B_2 = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right)$$

is independent.

The theorem above actually has another important consequence–it tells us that we do not need to check either of e_3 or e_4 for independence from the list B_2. Indeed, since A is invertible, the system

$$A\mathbf{x} = \mathbf{b}$$

is consistent for every 4×1 vector (i.e., any vector in $\mathbb{R}^4$). Thus B_2 is spanning in addition to independent, and forms a basis for $\mathbb{R}^4$.

Our final theorem in this section relates the idea of a direct sum of subspaces to the concept of a basis–indeed, we will see that the ideas are intricately intertwined. Recall that V is a *sum* of subspaces U and W, denoted

$$V = U + W,$$

if every vector in V can be written as a sum of a vector in U and a vector in W. The sum is *direct*, denoted

$$V = U \oplus W,$$

if

$$U \cap W = \{\mathbf{0}\}.$$

Theorem 2.34. Let U be a subspace of a finite dimensional vector space V. Then there is a subspace W of V so that

$$V = U \oplus W.$$

Proof. U is a (finite dimensional) vector space in its own right, and has a basis

$$S_u = (u_1,\ u_2,\ \ldots,\ u_n).$$

Clearly S_u is independent in V, and if $\text{span}\,(S_u) \neq V$, then using the previous theorem, we may extend S_u to a basis for V, say

$$S_v = (u_1,\ u_2,\ \ldots,\ u_n,\ w_1,\ \ldots,\ w_m).$$

Set $W = \text{span}\,(S_w)$, where

$$S_w = (w_1,\ \ldots,\ w_m).$$

We wish to show that W is a vector space.

Let $v_1,\ v_2 \in \text{span}\,(S_w)$, and $\alpha \in \mathbb{F}$; so all of $v_1,\ v_2,\ v_1 + v_2$, and αv_1 are linear combinations of vectors in S_w. Thus $v_1 + v_2 \in W$ and $\alpha v_1 \in W$, so that W is closed under addition and scalar multiplication and is a subspace of V.

Now since

$$S_v = (u_1,\ u_2,\ \ldots,\ u_n,\ w_1,\ \ldots,\ w_m) = S_u \cup S_w$$

is a basis for V, it is clear that every vector in V is a sum of a vector in U and a vector in W, that is

$$V = U + W.$$

It remains to show that the sum is direct; so consider $v \in U \cap W$. We may write v as a linear combination of vectors in U, as well as a linear combination of vectors in W, say

$$v = \alpha_1 u_1 + \ldots + \alpha_n u_n = \beta_1 w_1 + \ldots + \beta_m w_m.$$

This implies that

$$\alpha_1 u_1 + \ldots + \alpha_n u_n - \beta_1 w_1 - \ldots - \beta_m w_m = \mathbf{0};$$

but

$$S_v = (u_1,\ u_2,\ \ldots,\ u_n,\ w_1,\ \ldots,\ w_m)$$

is an independent list, so $\alpha_i = \beta_j = \mathbf{0}$, so that $v = \mathbf{0}$. Thus the sum is direct.

We record as a corollary an observation which is easily established using the technique in the proof:

Corollary. Let S_1 and S_2 be two finite lists of vectors from a (finite-dimensional) vector space V. If $S_1 \cup S_2$ is an independent list, then

$$\operatorname{span}(S_1) \cap \operatorname{span}(S_2) = \{\mathbf{0}\}.$$

Example. The space $\mathcal{U}_2(\mathbb{R})$ of 2×2 upper triangular matrices is a subspace of the vector space $\mathcal{M}_2(\mathbb{R})$ of 2×2 matrices. Find the subspace W guaranteed by the theorem so that

$$\mathcal{M}_2(\mathbb{R}) = \mathcal{U}_2(\mathbb{R}) \oplus W.$$

We can answer the question by comparing a standard basis for $\mathcal{U}_2(\mathbb{R})$ with a standard basis for $\mathcal{M}_2(\mathbb{R})$; we have seen that

$$S_u = \operatorname{span}\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

is a basis for $\mathcal{U}_2(\mathbb{R})$, and that

$$S_m = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

is a basis for $\mathcal{M}_2(\mathbb{R})$. Thus

$$W = \left\{ \begin{pmatrix} 0 & 0 \\ \alpha & 0 \end{pmatrix} \middle| \alpha \in \mathbb{R} \right\}$$

is the desired vector space.