Chapter 8: Bases and Dimension

8.1 Basis

Definition 8.1.1: Let V be a vector space. Then a subset $\mathscr{B}$ of V is said to be a *basis* for V if

1. $\mathscr{B}$ is linearly independent.
2. $\langle\mathscr{B}\rangle = V$, i.e., $\mathscr{B}$ spans V.

Remark: Most of the time V is a subspace of $\mathbb{R}^m$. Occasionally V is assumed to be a subspace of $M_{m,n}$ or P_n. It does not hurt to assume V is a subspace of $\mathbb{R}^m$.

Example 8.1.1: Let $V = \mathbb{R}^m$. Then $\mathscr{B} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ is a basis for V (recall that all the entries of $\boldsymbol{e}_i$ is zero, except the i-th entry being 1). It is called the *standard basis.*
Answer: Obviously $\mathscr{B}$ is linearly independent.
Also, for any $\alpha = (v_1, \ldots, v_m)^t \in V$, $\alpha = \sum_{i=1}^{m} v_i\boldsymbol{e}_i \in \langle\mathscr{B}\rangle$. So $\langle\mathscr{B}\rangle = V$. ■

Example 8.1.2: A vector space can have different bases. Example, $\mathscr{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ is a basis and $\mathscr{A} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$ is also a basis for $\mathbb{R}^2$.

Example 8.1 Math major only: $V = M_{2,2}$. Let

$$E^{1,1} = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, E^{1,2} = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix},$$

$$E^{2,1} = \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}, E^{2,2} = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix},$$

Then $\mathscr{B} = \{E^{1,1}, E^{1,2}, E^{2,1}, E^{2,2}\}$ is a basis for V.
Check:
Obviously $\mathscr{B}$ is linearly independent (exercise).
Also for any $A \in V$,

$$A = \begin{bmatrix}a & b\\c & d\end{bmatrix} = aE^{1,1} + bE^{1,2} + cE^{2,1} + dE^{2,2}.$$

So $\langle\mathscr{B}\rangle = M_{2,2}$.

Example 8.2 Math major only: Let $V = M_{m,n}$. For $1 \le i \le m$, $1 \le j \le n$, let $E^{i,j}$ be the $m \times n$ matrix with (i, j)-th entry equal to 1 and all other entries equal to 0.
Then $\{E^{i,j} \mid 1 \le i \le m,\ 1 \le j \le n\}$ is a basis for V. (exercise).

Example 8.3 Math major only: Let $V = P_n$.
Then $1, x, x^2, \ldots, x^n$ is a basis.
It is easy to show that $S = \{1, x, x^2, \ldots, x^n\}$ is linearly independent.
Also any polynomial

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

is a linear combinations of S.

Combining Theorems 7.2.8 and 7.3.5 we have

Theorem 8.1.2: *Suppose that A is a square matrix of order m. The columns of A form a basis for $\mathbb{R}^m$ if and only if A is nonsingular.*

From Theorem 7.3.5 we have

Theorem 8.1.3: *Let S be a finite subset of $\mathbb{R}^m$. Then basis for $\langle S\rangle$ exists. In fact, there exists a subset T of S such that T is a basis for $\langle S\rangle$.*

This theorem can be extended to any vector space, for example a subspace of $M_{m,n}$. Following is an extension.

Theorem 8.1.4: *Let $S = \{\alpha_1, \dots, \alpha_n\}$ be a finite subset of a vector space. Then basis for $\langle S\rangle$ exists. In fact, there exists a subset $\mathscr{B}$ of S such that $\mathscr{B}$ is a basis for $\langle S\rangle$.*

Before to prove the above theorem we show a useful lemma first.

Lemma 8.1.5: *Let S be a finite subset of a vector space. If $\alpha \in S$ is linearly dependent on other vectors in S, then $\langle S\rangle = \langle S \setminus \{\alpha\}\rangle$.*

Proof: It is clearly that $\langle S \setminus \{\alpha\}\rangle \subseteq \langle S\rangle$. So, we only need to show that $\langle S\rangle \subseteq \langle S \setminus \{\alpha\}\rangle$.

Proof of Theorem 8.1.4:

8.2 Dimension

Theorem 8.2.1 (Steintz Replacement Theorem)**:** *Let V be a vector space. Suppose $V = \langle \alpha_1, \dots, \alpha_n\rangle$. Then every linearly independent set $\{\beta_1, \dots, \beta_m\}$ contains at most n elements.*

Proof:

Corollary 8.2.2: *If a vector space has one basis with n elements, then all the other bases also have n elements.*

Proof: Suppose $\mathscr{A} = \{\alpha_1, \ldots, \alpha_n\}$ and $\mathscr{B} = \{\beta_1, \ldots, \beta_m\}$ are bases of a vector space. Since $V = \langle \mathscr{A} \rangle$ and $\mathscr{B}$ is linearly independent, by Theorem 8.2.1 $m \leq n$.

We change the role of $\mathscr{A}$ and $\mathscr{B}$, we will obtain that $n \leq m$.

Hence $m = n$. □

Definition 8.2.3: Let V be a nonzero vector space. Suppose $\{\alpha_1, \ldots, \alpha_t\}$ is a basis for V. Then t is called the *dimension* of V and is denoted by $t = \dim V$ and V is called a *finite dimensional vector space*. For convenience, we define $\dim\{\mathbf{0}\} = 0$.

Remark 8.2.4: By Corollary 8.2.2, the dimension is well-defined if a vector space contains a basis. So the next question is whether a vector space has a basis.

Corollary 8.2.5: *Suppose $m > n$. Then any m vectors in an n-dimensional vector space must be linearly dependent.*

Corollary 8.2.5 just follows from Theorem 8.2.1. We provide a directed proof for Corollary 8.2.5 as follows:

Proof: Suppose that $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis of the vector space V. Let $R = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$, where $m > n$. We will now construct a nontrivial relation of linear dependence on R.

Since $\langle S \rangle = V$, each $\boldsymbol{u}_i$ can be written as a linear combination of the vectors in S. This means there exist $a_{ij} \in \mathbb{R}$, $1 \leq i \leq n$, $1 \leq j \leq m$, such that

$$\begin{aligned}
\boldsymbol{u}_1 &= a_{11}\boldsymbol{v}_1 + a_{21}\boldsymbol{v}_2 + \cdots + a_{n1}\boldsymbol{v}_n = \sum_{i=1}^{n} a_{i1}\boldsymbol{v}_i \\
\boldsymbol{u}_2 &= a_{12}\boldsymbol{v}_1 + a_{22}\boldsymbol{v}_2 + \cdots + a_{n2}\boldsymbol{v}_n = \sum_{i=1}^{n} a_{i2}\boldsymbol{v}_i \\
\vdots \quad & \quad \vdots \\
\boldsymbol{u}_m &= a_{1m}\boldsymbol{v}_1 + a_{2m}\boldsymbol{v}_2 + \cdots + a_{nm}\boldsymbol{v}_n = \sum_{i=1}^{n} a_{im}\boldsymbol{v}_i
\end{aligned}$$

Now we form the homogeneous system of n equations in the m unkowns, $x_1, x_2, \ldots, x_m$, where the coefficients are the just-discovered scalars a_{ij},

$$\begin{aligned}
\sum_{j=1}^{m} a_{1j}x_j &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = 0 \\
\sum_{j=1}^{m} a_{2j}x_j &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = 0 \\
\vdots \quad & \quad \vdots \\
\sum_{j=1}^{m} a_{nj}x_j &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = 0
\end{aligned}$$

This is a homogeneous system with more unknowns than equations. So there are infinitely many solutions. Choose a nontrivial solution and denote it by $x_1 = c_1$, $x_2 = c_2$, $\ldots$, $x_m = c_m$. As a solution to the homogeneous system, we then have

$$\begin{aligned}
\sum_{j=1}^{m} a_{1j}c_j &= a_{11}c_1 + a_{12}c_2 + \cdots + a_{1m}c_m = 0 \\
\sum_{j=1}^{m} a_{2j}c_j &= a_{21}c_1 + a_{22}c_2 + \cdots + a_{2m}c_m = 0 \\
\vdots \quad & \quad \vdots \\
\sum_{j=1}^{m} a_{nj}c_j &= a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nm}c_m = 0
\end{aligned}$$

The scalars c_1, c_2, $\ldots$, c_m will provide the nontrivial relation of linear dependence we desire,

$$\begin{aligned}
c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_m\boldsymbol{u}_m &= \sum_{j}^{m} c_j\boldsymbol{u}_j \\
&= \sum_{j}^{m} c_j \left(\sum_{i}^{n} a_{ij}\boldsymbol{v}_i \right) = \sum_{j}^{m}\sum_{i}^{n} c_j a_{ij}\boldsymbol{v}_i = \sum_{i}^{n}\sum_{j}^{m} c_j a_{ij}\boldsymbol{v}_i \\
&= \sum_{i}^{n} \left(\sum_{j}^{m} a_{ij}c_j \right) \boldsymbol{v}_i = \sum_{i}^{n} 0\boldsymbol{v}_i = \mathbf{0}.
\end{aligned}$$

Hence R is linearly dependent. □

Example 8.4 Math major only: $\dim \mathbb{R}^m = m$.

Example 8.5 Math major only: $\dim M_{mn} = mn$. See Example 8.2.

Example 8.6 Math major only: $\dim P_n = n+1$. See Example 8.3.

Example 8.7 Math major only: Let S_2 be the set of 2×2 symmetric matrices. For $A \in S_2$,

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

We can show that

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

is a basis for S_2. Hence $\dim S_2 = 3$.

Example 8.8 Math major only: Let $\mathbb{R}[x]$ be the set of all real polynomials. As $\{1, x, x^2, x^3, \dots\}$ being linearly independent, so $\dim \mathbb{R}[x]$ does not exist (or we can write $\dim \mathbb{R}[x] = \infty$).

Lemma 8.2.6: *Let V be a vector space and $\alpha_1, \dots, \alpha_k, \alpha \in V$. Suppose $S = \{\alpha_1, \dots, \alpha_k\}$ is linearly independent and $\alpha \notin \langle S\rangle$. Then $S' = \{\alpha_1, \dots, \alpha_k, \alpha\}$ is linearly independent.*

Proof: Let the relation of linear dependence of S' be

$$a_1\alpha_1 + \cdots + a_k\alpha_k + a\alpha = \mathbf{0}.$$

Theorem 8.2.7: *In a finite dimensional vector space, any linearly independent set of vectors can be extended to a basis.*

Proof: Let $\{\beta_1, \dots, \beta_n\}$ be a linearly independent set in an m-dimensional vector space V. Let $\{\alpha_1, \dots, \alpha_m\}$ be a basis of V. Clearly $n \le m$ and $\{\beta_1, \dots, \beta_n, \alpha_1, \dots, \alpha_m\}$ spans V. If $n = 0$, then there is nothing to prove. So we assume $n > 0$. Thus $\{\beta_1, \dots, \beta_n, \alpha_1, \dots, \alpha_m\}$ is linearly dependent. Then there are $b_1, \dots, b_n, a_1, \dots, a_m \in \mathbb{R}$ not all zero such that $\sum\limits_{i=1}^{n} b_i\beta_i + \sum\limits_{j=1}^{m} a_j\alpha_j = \mathbf{0}$. We claim that at least one $a_j \neq 0$. For otherwise, if all the a_j's are zero, then we have $\sum\limits_{i=1}^{n} b_i\beta_i = \mathbf{0}$ and by the assumption, $b_1 = \cdots = b_n = 0$. This is impossible.

Thus by Lemma 8.2.6 $\{\beta_1, \dots, \beta_n, \alpha_1, \dots, \alpha_{j-1},\ \alpha_{j+1}, \dots \alpha_m\}$ still spans V. If $n > 1$, then this set is linearly dependent and we can apply the above argument to discard another α_j and still obtain a spanning set of V. We continue this process until we get m spanning vectors, n of which are $\beta_1, \dots, \beta_n$. This is a required basis. $\square$

Remark 8.2.8: From the proof above, we see that there are more than one way of extending a linearly independent set to a basis.

Let V be a finite dimensional vector space and let W be a subspace of V. What is the dimension of W? That means whether W contains a basis. Is there any relation between the dimension of W and the dimension of V? We shall answer these questions below.

Theorem 8.2.9: *A subspace W of an m-dimensional vector space V is a finite dimensional vector space of dimension at most m.*

Proof: If $W = \{\mathbf{0}\}$, then W is 0-dimensional.

Corollary 8.2.10: *Let V be a subspace of $\mathbb{R}^m$. There exists a basis for V.*

Corollary 8.2.11: *Let $S = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{R}^m$. Then $\dim \langle S \rangle \leq n$ and $\dim \langle S \rangle \leq m$.*

Theorem 8.2.12: *Let W be a subspace of V with a basis $\mathscr{B} = \{\alpha_1, \ldots, \alpha_n\}$. Assume that $\dim V = m$. Then there exists a basis $\mathscr{B} \cup \{\alpha_{n+1}, \ldots, \alpha_m\}$ of V for some vectors $\alpha_{n+1}, \ldots, \alpha_m$ in V.*

Proof: This follows from Theorem 8.2.7. $\square$

Remark 8.2.13: Every infinite dimensional vector space also has a basis. However to show this, we have to require axiom of choice or apply Kuratowski-Zorn's lemma, which is beyond the scope of this course.

Theorem 8.2.14: *Let V be an n-dimensional vector space and $\mathscr{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a set of vectors in V. Then the following statements are equivalent:*

(a) $\mathscr{A}$ is a basis.

(b) $\mathscr{A}$ is linearly independent.

(c) $V = \langle \mathscr{A} \rangle$.

Proof:

(a)$\Rightarrow$(b): Clear.

Corollary 8.2.15: *Suppose W_1 and W_2 are two subspaces of V. If $W_1 \subseteq W_2$ and $\dim W_1 = \dim W_2 < \infty$, then $W_1 = W_2$.*

Proof: Suppose $\{\alpha_1, \dots, \alpha_m\}$ is a basis of W_1. Then $\{\alpha_1, \dots, \alpha_m\} \subset W_2$ is linearly independent. Since $\dim W_1 = \dim W_2$, by Theorem 8.2.14 it is also a basis of W_2. Therefore, $W_1 = W_2$. □

Remark 8.2.16: The condition $W_1 \subseteq W_2$ is crucial. For taking $W_1 = \langle (1,0)^t \rangle$ and $W_2 = \langle (0,1)^t \rangle$, it is easy to see that $W_1 \neq W_2$ yet $\dim W_1 = \dim W_2 = 1$.

8.3 Ranks and Nullity of a Matrix

Since the RREF of a matrix A is unique, the number of non-zero rows of the RREF of A is denoted by $r(A)$, which is called the *rank* of A (has already been defined in Chapter 5).

Definition 8.3.1: Suppose that $A \in M_{m,n}$.

1. The *nullity* of A is the dimension of the null space of A, i.e., $n(A) = \dim(\mathcal{N}(A))$.
2. The *column rank* of A is the dimension of the column space of A, $\operatorname{colrank}(A) = \dim(\mathcal{C}(A))$.
3. The *row rank* of A is the dimension of the row space of A, $\operatorname{rowrank}(A) = \dim(\mathcal{R}(A))$.

By Theorem 7.3.5, we have

Theorem 8.3.2: *Suppose that $A \in M_{m,n}$. Then $r(A) = \operatorname{colrank}(A)$.*

In other sections of MATH1030, $r(A)$ is defined to be $\operatorname{colrank}(A)$ directly.

Example 8.3.1: Let us compute the rank and nullity of

$$A = \begin{pmatrix} 2 & -4 & -1 & 3 & 2 & 1 & -4 \\ 1 & -2 & 0 & 0 & 4 & 0 & 1 \\ -2 & 4 & 1 & 0 & -5 & -4 & -8 \\ 1 & -2 & 1 & 1 & 6 & 1 & -3 \\ 2 & -4 & -1 & 1 & 4 & -2 & -1 \\ -1 & 2 & 3 & -1 & 6 & 3 & -1 \end{pmatrix}.$$

We have

$$\operatorname{rref}(A) = \begin{pmatrix} \textcircled{1} & -2 & 0 & 0 & 4 & 0 & 1 \\ 0 & 0 & \textcircled{1} & 0 & 3 & 0 & -2 \\ 0 & 0 & 0 & \textcircled{1} & -1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

From $\operatorname{rref}(A)$ we record $D = \{1, 3, 4, 6\}$ and $F = \{2, 5, 7\}$.

By Theorem 7.2.10, $\{A_{*1}, A_{*3}, A_{*4}, A_{*6}\}$ is a basis of $\mathcal{C}(A)$. So $r(A) = \operatorname{colrank}(A) = 4$.

By Theorem 7.2.9, $\{(2,1,0,0,0,0,0)^t, (-4,0,-3,1,1,0,0)^t, (-1,0,2,3,0,-1,1)^t\}$ is a basis of $\mathcal{N}(A)$. Hence $n(A) = 3$.

Now we have $r(A) + n(A) = 4 + 3 = 7 =$ the number of column of A. ■

Theorems 7.2.9 and 7.2.10 show that

Theorem 8.3.3 (Dimension Formula)**:** *Suppose $A \in M_{m,n}$. Then*

$$r(A) + n(A) = n.$$

Corollary 8.3.4: *Let A be an $m \times n$ matrix. Then*

$$r(A) = r(A^t).$$

Equivalently

$$\dim \mathcal{C}(A) = \dim \mathcal{R}(A).$$

Proof:

Corollary 8.3.5: *Let A be an $m \times n$ matrix. Then*

$$r(A) = \text{rowrank}(A).$$

Theorem 8.3.6: *Suppose that $A \in M_n$. The following are equivalent.*

1. *A is nonsingular.*
2. *$r(A) = n$.*
3. *$n(A) = 0$.*

Proof:

With a new equivalence for a nonsingular matrix, we can update Theorem 7.2.8 which becomes a list requiring double digits to number.

Theorem 8.3.7: *Suppose that $A \in M_n$. The following are equivalent.*

1. *A is nonsingular.*
2. *A is row equivalent to I_n.*
3. *$\mathcal{N}(A) = \{\mathbf{0}_n\}$.*
4. *The linear system $\mathcal{LS}(A, \boldsymbol{b})$ has a unique solution for every possible choice of $\boldsymbol{b}$.*
5. *A is invertible.* Skip it if Chapter 5 has not been taught.

6. *The columns of A form a linearly independent set.*

7. *The column space of A is $\mathbb{R}^n$, i.e., $\mathcal{C}(A) = \mathbb{R}^n$.*

8. *The columns of A form a basis for $\mathbb{R}^n$.*

9. *The rank of A is n, i.e., $r(A) = n$.*

10. *The nullity of A is zero, i.e., $n(A) = 0$.*