Linear Algebra & Differential Equations

Elif Uskuplu

If you see any mistake, please email me (euskuplu@usc.edu).

8 Week 8

8.1 Basis of a vector space

The concept of a *basis* in a vector space is akin to the *building blocks* or *foundation* essential for constructing any vector within that space. Imagine these basis vectors as the fundamental pieces from which all vectors in the space can be composed, much like LEGO bricks can be combined to build various structures.

A basis acts as the **fundamental set of blocks** or bricks, enabling the construction of any vector in the vector space, similar to how a diverse set of LEGO pieces allows for the creation of myriad designs.

Just as each building has a unique arrangement of blocks, every vector in a vector space can be **uniquely represented** as a combination of the basis vectors. This uniqueness ensures that no two vectors are identical when they are composed of the basis vectors in different proportions.

The essence of a good set of building blocks is its **minimality and sufficiency**—having exactly what is needed, no more, no less. Correspondingly, a basis for a vector space is the *smallest set of vectors* that can span the entire space, emphasizing the efficiency and necessity of each basis vector.

A basis provides a **versatile and adaptable foundation** for the vector space, akin to selecting the right foundational elements for constructing buildings of various purposes. The choice of basis can simplify the representation and manipulation of vectors, much like choosing the appropriate building blocks can facilitate construction.

Consider a basis as establishing a **coordinate system** within the vector space. Similar to how a grid of streets and addresses allows for the precise location of any place within a city, the basis vectors enable the precise identification of any vector in the space through a unique set of coordinates.

In summary, a basis of a vector space is the *groundwork* or *substructure*, providing a unique, minimal, and sufficient foundation from which any vector in the space can be constructed. It offers a versatile framework that underpins the entire space, facilitating the creation of a coordinate system to precisely locate and represent any vector within that domain.

Definition 8.1. *A **basis** of a vector space V over a field F is defined as a set of vectors $\mathbf{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in V that satisfies two main criteria:*

1. *The vectors in $\mathbf{B}$ are linearly independent.*
2. *The set $\mathbf{B}$ spans V.*

If a set $\mathbf{B}$ meets these conditions, then it is a basis for the vector space V, and the dimension of V is defined as the number of vectors in $\mathbf{B}$, denoted by $dim(V)$.

Example. The standard basis for the n-dimensional real vector space $\mathbb{R}^n$ consists of n vectors, where each vector has exactly one component equal to 1 and all other components equal to 0. These basis vectors are denoted as follows:

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, \ldots, 0), \quad \ldots, \quad \mathbf{e}_n = (0, 0, \ldots, 1)$$

As an exercise, you can prove that these are linearly independent and spanning $\mathbb{R}^n$. Also, we conclude $dim(\mathbb{R}^n) = n$.

Example. The space $M_n(\mathbb{R})$ of all $n \times n$ matrices over the real numbers is a vector space. The standard basis for this vector space consists of n^2 matrices, where each matrix has a single entry of 1 in a unique position and all other entries are 0. These basis matrices can be denoted as E_{ij}, where $1 \leq i, j \leq n$, and the matrix E_{ij} has a 1 in the i-th row and j-th column, and 0s elsewhere.

$$E_{ij} = (e_{kl}) \quad \text{where} \quad e_{kl} = \begin{cases} 1 & \text{if } k = i \text{ and } l = j, \\ 0 & \text{otherwise.} \end{cases}$$

For example, in the case of 2×2 matrices, the standard basis for $M_2(\mathbb{R})$ consists of the following matrices:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

As an exercise, you can prove that E_{ij}s are linearly independent and spanning $M_n(\mathbb{R})$. Also, we conclude $dim(M_n(\mathbb{R})) = n^2$.

Example. The vector space $M_{m\times n}(\mathbb{R})$ consists of all $m \times n$ matrices with real number entries. The standard basis for this space is a set of matrices where each matrix has one entry of 1 in a distinct position and 0 in all other positions. These basis elements are denoted as E_{ij}, where $1 \leq i \leq m$ and $1 \leq j \leq n$. The matrix E_{ij} has a 1 in the i-th row and j-th column and 0s elsewhere.

$$E_{ij} = (e_{kl}) \quad \text{for} \quad 1 \leq k \leq m, \ 1 \leq l \leq n, \quad \text{where} \quad e_{kl} = \begin{cases} 1 & \text{if } k = i \text{ and } l = j, \\ 0 & \text{otherwise.} \end{cases}$$

For instance, in the case of 3×2 matrices, part of the standard basis for $M_{3\times 2}(\mathbb{R})$ includes matrices such as:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \ldots$$

As an exercise, you can prove that E_{ij}s are linearly independent and spanning $M_{m\times n}(\mathbb{R})$. Also, we conclude $dim(M_{m\times n}(\mathbb{R})) = mn$.

Example. The vector space $P_n(\mathbb{R})$ consists of all polynomials with real coefficients and degree at most n. A basis for this vector space is a set of polynomials that are linearly independent and span the entire space. The standard basis for $P_n(\mathbb{R})$ is given by the polynomials:

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

This means that any polynomial $p(x) \in P_n(\mathbb{R})$ of degree at most n can be written uniquely as a linear combination of these basis polynomials. Specifically, a polynomial $p(x)$ can be expressed as:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

where $a_0, a_1, \ldots, a_n$ are real numbers (coefficients). Also, we conclude $dim(P_n(\mathbb{R})) = n+1$.

Remark. If V is the trivial vector space $\{\mathbf{0}\}$, then we define its dimension to be zero, namely, $dim(\{\mathbf{0}\}) = 0$.

Unusual example for curious students :) The vector space Fun$(\mathbb{R}, \mathbb{R})$ encompasses all functions mapping real numbers to real numbers. Unlike finite-dimensional vector spaces, or even some infinite-dimensional spaces with more structure, Fun$(\mathbb{R}, \mathbb{R})$ poses significant challenges in identifying a basis.

A *Hamel basis* for Fun$(\mathbb{R}, \mathbb{R})$ would allow every function within the space to be uniquely expressed as a finite linear combination of basis functions. However, explicitly constructing or even identifying a Hamel basis for such a general and vast space is impractical with conventional mathematical approaches.

The existence of a Hamel basis in Fun$(\mathbb{R}, \mathbb{R})$, as in any vector space, relies on the Axiom of Choice. This foundational principle in set theory enables the assertion of a basis's existence without the necessity for its explicit construction, especially in spaces as large as Fun$(\mathbb{R}, \mathbb{R})$.

While the Axiom of Choice facilitates the theoretical foundation for the existence of a basis in Fun$(\mathbb{R}, \mathbb{R})$, the practical implications of such a basis are limited. In real-world applications, mathematicians often resort to alternative frameworks better suited to the nature of the functions and analyses in question.

Although the textbook includes full proofs of the following theorems and others like them, it's worth emphasizing their significance due to their utility.

Remark. A basis of a vector space V can be considered as the minimum spanning set and the maximum linearly independent set.

Theorem 8.1. *If $dim(V) = n$, then any set of n linearly independent vectors in V is a basis for V.*

Proof. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a set of n linearly independent vectors in V. Since $\dim(V) = n$, it follows that a basis for V must consist of exactly n-many vectors. We know that a set of vectors forms a basis for V if and only if it is linearly independent and spans V.

Since S is linearly independent by assumption, it remains to show that S spans V. Assume, for the sake of contradiction, that S does not span V. Then, there exists a vector $\mathbf{v} \in V$ that cannot be expressed as a linear combination of vectors in S. Adding $\mathbf{v}$ to S creates a set $S' = S \cup \{\mathbf{v}\}$ of $(n+1)$-many vectors. Since $\mathbf{v}$ cannot be written as a linear combination of the vectors in S, the set S' is linearly independent. This contradicts the fact that the maximum number of linearly independent vectors in V is n, as $\dim(V) = n$.

Therefore, our assumption that S does not span V must be false. Hence, S is linearly independent and spans V, making it a basis for V. □

Example. Consider the vector space $M_{2\times 2}(\mathbb{R})$ of 2×2 real matrices. The dimension of this space is $n = 4$

Consider the following set of 2×2 matrices:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

If $a, b, c, d \in \mathbb{R}$ such that $aA_1 + bA_2 + cA_3 + dA_4 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Then, we get

$$\begin{aligned} a + b &= 0 \\ c + d &= 0 \\ c - d &= 0 \\ a - b &= 0 \end{aligned}$$

It is easy to see that $a = b = c = d = 0$, so $\{A_1, A_2, A_3, A_4\}$ is linearly independent. By the theorem, this is a basis for $M_2(\mathbb{R})$.

Theorem 8.2. *If $dim(V) = n$, then any set of n vectors in V that spans V is a basis for V.*

Proof. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a set of n vectors in V that spans V. Since $\dim(V) = n$, the dimension of V indicates that the maximum number of linearly independent vectors in V is n. To prove that S is a basis for V, we need to show that it is also linearly independent.

Assume for the sake of contradiction that S is not linearly independent. This means there exists a non-trivial linear combination of the vectors in S that equals the zero vector, i.e., there exist scalars $a_1, a_2, \ldots, a_n$, not all zero, such that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n = \mathbf{0}$.

However, since S spans V, any vector in V can be expressed as a linear combination of vectors in S. If S were not linearly independent, we could remove at least one vector from S and still span V, contradicting the fact that $\dim(V) = n$, namely, the minimal spanning set contains n vectors. This contradiction implies that our assumption is false, and thus S must be linearly independent.

Therefore, since S spans V and is linearly independent, S is a basis for V. □

Example. To show that the polynomials 3, $1 + x$, and $2 + x^2$ span $P_2(\mathbb{R})$, we must demonstrate that any polynomial $p(x)$ of degree at most 2 can be written as a linear combination of these polynomials.

Let $p(x) = ax^2 + bx + c$ be an arbitrary element of $P_2(\mathbb{R})$. We want to find scalars k_1, k_2, k_3 such that $k_1 \cdot 3 + k_2 \cdot (1 + x) + k_3 \cdot (2 + x^2) = ax^2 + bx + c$.

By equating coefficients, we obtain the system of equations:

$$\begin{aligned} 3k_1 + k_2 + 2k_3 &= c, \\ k_2 &= b, \\ k_3 &= a. \end{aligned}$$

This system of equations has a solution for k_1, k_2, k_3 for any given a, b, c in $\mathbb{R}$, namely:

$$\begin{aligned} k_1 &= \frac{1}{3}(c - b - 2a), \\ k_2 &= b, \\ k_3 &= a. \end{aligned}$$

Thus, any polynomial $p(x) = ax^2 + bx + c$ in $P_2(\mathbb{R})$ can be written as a linear combination of 3, $1+x$, and $2+x^2$. Therefore, these polynomials span $P_2(\mathbb{R})$. Since $dim(P_2(\mathbb{R})) = 3$, by the theorem, $\{3, 1 + x, 2 + x^2\}$ is a basis for $P_2(\mathbb{R})$.

Now, we can combine these as a one corollary:

Corollary 8.1. *If $dim(V) = n$ and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of n vectors in V, the following statements are equivalent:*

1. *S is a basis for V.*
2. *S is linearly independent.*
3. *S spans V.*

Vector Space	Standard Basis	Dimension
$\mathbb{R}^n$	$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$	n
$M_{m\times n}(\mathbb{R})$	$\{E_{ij} \mid 1 \le i \le m, 1 \le j \le n\}$	mn
$M_n(\mathbb{R})$ (Square Matrices)	$\{E_{ij} \mid 1 \le i, j \le n\}$	n^2
$P_n(\mathbb{R})$	$\{1, x, x^2, \ldots, x^n\}$	$n+1$
$\mathbb{C}^n$	$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$	n

Remark. Suppose that V is a vector space with $\dim[V] = n$, and let $S = \{v_1, v_2, \ldots, v_k\}$ be a subset of V.

	$k < n$	$k > n$	$k = n$
S is linearly independent?	Maybe	No	Maybe
S spans V?	No	Maybe	Maybe
S is a basis?	No	No	Maybe

Another important observation is the following corollary:

Corollary 8.2. *Let S be a subspace of finite dimensional vector space V. If $dim(V) = n$, then*

$$dim(S) \le n.$$

Furthermore, if $dim(S) = n$, then $S = V$.

Proof. Assume for the sake of contradiction that $\dim(S) > n$. This would mean that there exists a set of $m > n$ linearly independent vectors in S. However, since $S \subseteq V$ and $\dim(V) = n$, there cannot be more than n linearly independent vectors in V, let alone in S. This contradiction implies that our assumption is false, and therefore $\dim(S) \le n$.

Now assume that $\dim(S) = n$. Then there exists a basis for S consisting of n vectors. Since V also has a dimension of n, this basis for S must also be a basis for V, as it is a set of n linearly independent vectors that spans a subspace of V. Hence, every vector in V can be expressed as a linear combination of vectors in S, which means $S = V$. □

Example. We want to find the dimension of the null space of the matrix

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

To do this, we solve the homogeneous system $A\mathbf{x} = \mathbf{0}$.

First, perform row reduction on the augmented matrix $[A|\mathbf{0}]$:

$$\left[\begin{array}{cc|c} 1 & 3 & 0 \\ -2 & -6 & 0 \end{array}\right]$$

After performing the row operations, we get:

$$\left[\begin{array}{cc|c} 1 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Then we get y as free variable and $x = -3y$, so

$$nullspace(A) = \{(-3y, 1y) \mid y \in \mathbb{R}\} = span\{(-3, 1)\}.$$

Since a single nonzero vector is linearly independent, $\{(-3, 1)\}$ is a basis for $nullspace(A)$. Thus $dim(nullspace(A)) = 1$.

Example. Let S be the subspace of $M_2(\mathbb{R})$ consisting of all upper triangular matrices. The general form of an upper triangular matrix in $M_2(\mathbb{R})$ is

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix},$$

where a, b, and c are real numbers.

A basis for S can be formed by the following matrices:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Each matrix in this set is linearly independent from the others, and any upper triangular matrix in $M_2(\mathbb{R})$ can be expressed as a linear combination of these matrices.

Therefore, the dimension of S, denoted as $\dim(S)$, is equal to the number of matrices in the basis for S, which is 3.

8.2 Change of Basis

We start with an important property of bases:

Theorem 8.3. *If V is a vector space with basis $\{v_1, v_2, \ldots, v_n\}$, then every vector $v \in V$ can be written uniquely as a linear combination of $v_1, v_2, \ldots, v_n$.*

Proof. Since $v_1, v_2, \ldots, v_n$ span V, every vector $v \in V$ can be expressed as

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n \quad (1)$$

for some scalars $a_1, a_2, \ldots, a_n$. Suppose also that

$$v = b_1v_1 + b_2v_2 + \cdots + b_nv_n \quad (2)$$

for some scalars $b_1, b_2, \ldots, b_n$. We will show that $a_i = b_i$ for each i, which will prove the uniqueness assertion of this theorem. Subtracting Equation (2) from Equation (1) yields

$$(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n = 0 \quad (3)$$

But $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, and so Equation (3) implies that

$$a_1 - b_1 = 0, \quad a_2 - b_2 = 0, \quad \ldots, \quad a_n - b_n = 0.$$

That is, $a_i = b_i$ for each $i = 1, 2, \ldots, n$. □

An **ordered basis** for a vector space V over a field F is a sequence (rather than a set) of linearly independent vectors in V that spans the entire vector space. Formally, an ordered basis is a finite sequence $(v_1, v_2, \ldots, v_n)$ such that for every vector $v \in V$, there exists a unique sequence of scalars $(a_1, a_2, \ldots, a_n) \in F$ for which

$$v = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n.$$

The main difference between an ordered basis and a simple basis is the emphasis on the sequence or order of the vectors, which influences the representation of vectors and linear operators in the context of the chosen basis.

Definition 8.2. *If $B = \{v_1, v_2, \ldots, v_n\}$ is an ordered basis for V and v is a vector in V, then the scalars $c_1, c_2, \ldots, c_n$ in the unique n-tuple $(c_1, c_2, \ldots, c_n)$ such that*

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

are called ***the components of** v **relative to the ordered basis** $B = \{v_1, v_2, \ldots, v_n\}$. We denote the column vector consisting of the components of v relative to the ordered basis B by $[v]_B$ and we call $[v]_B$ **the component vector of** v **relative to** B.*

Examples.

1. Consider the vector space $\mathbb{R}^2$ and an ordered basis $B = \{(1, 2), (3, 1)\}$. For a vector $v = (7, 5)$, the component vector of v relative to B can be represented as:

$$[v]_B = \begin{bmatrix} 8/5 \\ 9/5 \end{bmatrix}$$

2. In the vector space of polynomials of degree at most 2, $P_2(\mathbb{R})$, with an ordered basis $B = \{1, x, x^2\}$, consider a polynomial $p(x) = 4 + 3x + 2x^2$. The component vector of $p(x)$ relative to B is:

$$[p(x)]_B = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$$

This represents the coefficients of $p(x)$ relative to the basis B.

In the same vector space, if we consider another ordered basis $C = \{3, 1+x, 2+x^2\}$, consider the same polynomial $p(x) = 4 + 3x + 2x^2$. We already solved the spanning relation in the previous section. The component vector of $p(x)$ relative to C is:

$$[p(x)]_C = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}$$

3. For the vector space of 2×2 matrices, $M_{2\times 2}(\mathbb{R})$, with an ordered basis

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

for a matrix $A = \begin{pmatrix} 5 & 7 \\ 8 & 9 \end{pmatrix}$, the component vector of A relative to B is:

$$[A]_B = \begin{bmatrix} 5 \\ 7 \\ 8 \\ 9 \end{bmatrix}$$

This shows how the matrix A is decomposed into the components relative to the basis B.

Lemma 8.4. *Let V be a vector space with basis $B = \{v_1, v_2, \ldots, v_n\}$, let x and y be vectors in V, and let c be a scalar. Then we have:*

(a) $[x + y]_B = [x]_B + [y]_B$,

(b) $[cx]_B = c[x]_B$.

Proof. Easy, left as an exercise. □

Let V be a vector space, and let $B = \{b_1, b_2, \ldots, b_n\}$ and $C = \{c_1, c_2, \ldots, c_n\}$ be two bases for V. The **change-of-basis matrix from** B **to** C, denoted by $P_{C\leftarrow B}$, is the matrix that, when applied to the coordinate vector of a vector v with respect to B, yields the coordinate vector of v with respect to C. Formally, if $[v]_B$ is the coordinate vector of v relative to B, then $[v]_C = P_{C\leftarrow B}[v]_B$, where $[v]_C$ is the coordinate vector of v relative to C.

To construct $P_{C\leftarrow B}$, express each vector b_i of the basis B in terms of the basis C, and use these expressions to form the columns of $P_{C\leftarrow B}$. Specifically, if $b_i = a_{1i}c_1 + a_{2i}c_2 + \ldots + a_{ni}c_n$ in terms of C, then the i-th column of $P_{C\leftarrow B}$ is $\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$. In other words, we have

$$P_{C\leftarrow B} = \Big[[b_1]_C \; [b_2]_C \; \ldots \; [b_n]_C \Big]$$

Examples.

1. Let's consider a 2-dimensional vector space V and define two bases for V:

   - Basis $B = \{b_1, b_2\}$ where $b_1 = (1, 0)$ and $b_2 = (0, 1)$. This is the standard basis for V.
   - Basis $C = \{c_1, c_2\}$ where $c_1 = (2, 1)$ and $c_2 = (1, 1)$.

To construct the change-of-basis matrix from B to C, denoted $P_{C \leftarrow B}$, we need to express each vector b_i of the basis B in terms of the basis C.

Express b_1 in terms of C, solving for coefficients a_{11} and a_{21} such that $b_1 = a_{11}c_1 + a_{21}c_2$.

Express b_2 in terms of C, solving for coefficients a_{12} and a_{22} such that $b_2 = a_{12}c_1 + a_{22}c_2$.

These coefficients a_{ij} will form the columns of $P_{C \leftarrow B}$, where the i-th column is:

$$\begin{bmatrix} a_{1i} \\ a_{2i} \end{bmatrix}$$

After solving the equations, we find:

$$\begin{aligned} a_{11} &= 1, \quad & a_{21} &= -1, \\ a_{12} &= -1, \quad & a_{22} &= 2. \end{aligned}$$

Therefore, the change-of-basis matrix from B to C, $P_{C \leftarrow B}$, is:

$$P_{C \leftarrow B} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

This matrix, when applied to the coordinate vector of a vector v with respect to B, yields the coordinate vector of v with respect to C.

2. Let's consider $P_2(\mathbb{R})$ and define two bases:

- Basis $B = \{1, x, x^2\}$

- Basis $C = \{2 + x, x + x^2, 5 + x^2\}$

To find the change-of-basis matrix from B to C, denoted $P_{C \leftarrow B}$, we need to express each vector in B as a linear combination of vectors in C. That is, for each $b_i \in B$, we find coefficients a_{ij} such that

$$b_i = a_{i1}(2 + x) + a_{i2}(x + x^2) + a_{i3}(5 + x^2)$$

Specifically, we solve for:

1. $1 = a_{11}(2 + x) + a_{12}(x + x^2) + a_{13}(5 + x^2)$

2. $x = a_{14}(2 + x) + a_{15}(x + x^2) + a_{16}(5 + x^2)$

3. $x^2 = a_{17}(2 + x) + a_{18}(x + x^2) + a_{19}(5 + x^2)$

By matching coefficients of like terms (1, x, and x^2) on both sides of these equations, we obtain the coefficients that form the change-of-basis matrix:

$$P_{C \leftarrow B} = \begin{bmatrix} \frac{1}{7} & \frac{5}{7} & -\frac{5}{7} \\ -\frac{1}{7} & \frac{2}{7} & \frac{5}{7} \\ \frac{1}{7} & -\frac{2}{7} & \frac{2}{7} \end{bmatrix}$$

3. Consider the vector space $M_2(\mathbb{R})$, the space of 2×2 matrices with real number entries. We define two bases for $M_2(\mathbb{R})$:

   - The standard basis B:

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

   - Another basis C:

$$C = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \right\}$$

   To find the change-of-basis matrix $P_{C \leftarrow B}$ from B to C, we express each matrix in B as a linear combination of matrices in C. This involves solving for the coefficients that express the basis elements of B in terms of the basis elements of C.

   After solving for these coefficients, we obtain the change-of-basis matrix $P_{C \leftarrow B}$ as:

$$P_{C \leftarrow B} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

   This matrix, when applied to the coordinate vector of a matrix in B, yields the coordinate vector of that matrix in C.

Remark. Given a vector space V and two bases B and C for V, the change-of-basis matrix from B to C, denoted $P_{C \leftarrow B}$, transforms coordinate vectors relative to B into coordinate vectors relative to C. Similarly, the change-of-basis matrix from C to B, denoted $P_{B \leftarrow C}$, performs the inverse operation. This proof demonstrates that $P_{C \leftarrow B}$ and $P_{B \leftarrow C}$ are indeed inverses of each other.

Proof. Let $[v]_B$ be the coordinate vector of a vector $v \in V$ relative to the basis B, and let $[v]_C$ be the coordinate vector of v relative to the basis C. By definition, we have:

$$[v]_C = P_{C \leftarrow B}[v]_B \tag{1}$$

and

$$[v]_B = P_{B \leftarrow C}[v]_C \tag{2}$$

Substituting Equation (1) into Equation (2), we get:

$$\begin{aligned} [v]_B &= P_{B \leftarrow C}(P_{C \leftarrow B}[v]_B) \\ &= (P_{B \leftarrow C}P_{C \leftarrow B})[v]_B \end{aligned}$$

For this to be true for all $v \in V$, it must be that:

$$P_{B \leftarrow C}P_{C \leftarrow B} = I \tag{3}$$

where I is the identity matrix of appropriate size.

Similarly, by applying the change-of-basis matrices in the reverse order and following a similar argument, we can show that:

$$P_{C\leftarrow B}P_{B\leftarrow C} = I \tag{4}$$

Equations (3) and (4) together imply that $P_{C\leftarrow B}$ and $P_{B\leftarrow C}$ are inverses of each other. □

The following is another property of change of basis matrices.

Theorem 8.5. *Let V be a vector space with ordered bases A, B, and C. Then the change-of-basis matrix from A to C, denoted by $P_{C\leftarrow A}$, can be expressed as the product of the change-of-basis matrix from B to C, $P_{C\leftarrow B}$, and the change-of-basis matrix from A to B, $P_{B\leftarrow A}$. Formally, we have:*

$$P_{C\leftarrow A} = P_{C\leftarrow B}P_{B\leftarrow A}. \tag{5}$$

Proof. Given any vector $v \in V$, let $[v]_A$, $[v]_B$, and $[v]_C$ denote the coordinate vectors of v relative to bases A, B, and C, respectively. By the definition of a change-of-basis matrix, we have:

$$[v]_B = P_{B\leftarrow A}[v]_A, \tag{6}$$
$$[v]_C = P_{C\leftarrow B}[v]_B. \tag{7}$$

Substituting (1) into (2), we get:

$$\begin{aligned} [v]_C &= P_{C\leftarrow B}(P_{B\leftarrow A}[v]_A) \\ &= (P_{C\leftarrow B}P_{B\leftarrow A})[v]_A. \end{aligned}$$

Since $[v]_C = P_{C\leftarrow A}[v]_A$ by definition, it follows that:

$$P_{C\leftarrow A} = P_{C\leftarrow B}P_{B\leftarrow A}. \tag{8}$$

□

8.3 Row Space & Column Space

- The **row space** of a matrix A, denoted as rowspace(A), is the set of all possible linear combinations of its row vectors. Formally, if A is a matrix with rows $r_1, r_2, \ldots, r_m$, then:

  $$\text{rowspace}(A) = \text{span}\{r_1, r_2, \ldots, r_m\}.$$

  To find a basis for the row space of A, perform row reduction on A to obtain its row echelon form (REF) or reduced row echelon form (RREF). The non-zero rows in the REF or RREF form a basis for the row space of A.

- The **column space** of a matrix A, denoted as columnspace(A), is the set of all possible linear combinations of its column vectors. Formally, if A is a matrix with columns $c_1, c_2, \ldots, c_n$, then:

$$\text{columnspace}(A) = \text{span}\{c_1, c_2, \ldots, c_n\}.$$

  To find a basis for the column space of A, identify the columns with leading 1s in the REF or RREF of A. The corresponding columns in the original matrix A form a basis for the column space of A.

- The **nullity** of a matrix A is defined as the dimension of the null space of A. The null space of A, denoted as nullspace(A), is the set of all vectors x such that $Ax = \mathbf{0}$, where **0** is the zero vector.

  Formally, if A is an $m \times n$ matrix, then the nullity of A is given by:

$$\text{nullity}(A) = \dim(\text{nullspace}(A))$$

Remark 1. For any $m \times n$ matrix A, we have

$$\text{rowspace}(A^T) = \text{columnspace}(A).$$

Remark 2. The dimension of the row space (or rank) of A equals the dimension of the column space (or rank) of A. This is written as:

$$\dim(\text{rowspace}(A)) = \dim(\text{columnspace}(A)) = \text{rank}(A).$$

Example. Given the matrix B:

$$B = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 3 & 6 & -3 & 5 \\ 1 & 2 & -1 & -1 \\ 5 & 10 & -5 & 7 \end{pmatrix}$$

The RREF of matrix B is:

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

The row space of B can be represented by the non-zero rows obtained after converting B into its reduced row echelon form (RREF). The basis for the row space of B is:

$$\left\{ \begin{pmatrix} 1 & 2 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \right\}$$

indicating that the dimension of the row space is 2.

The column space of B is spanned by the columns of B corresponding to the leading one positions in its RREF. The basis for the column space of B is:

$$\left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ -1 \\ 7 \end{pmatrix} \right\}$$

This also indicates that the dimension of the column space is 2.

Now, we have very important result indicating a relationship between rank and nullity:

Theorem 8.6. *Let A be an $m \times n$ matrix. Then the rank of A plus the nullity of A is equal to n, the number of columns in A. Formally, this can be expressed as:*

$$\textit{rank}(A) + \textit{nullity}(A) = n$$

Proof. Let A be an $m \times n$ matrix. First, we reduce A to its RREF, denoted as A'. In A', the number of pivot columns (i.e., columns containing a leading 1) equals the rank of A, because these columns are linearly independent by construction.

The basis for the null space of A can be found by solving $Ax = 0$. Each free variable in the system corresponds to one basis vector in the null space. Therefore, the number of free variables, which equals the number of columns n minus the number of pivot columns (the rank of A), gives the dimension of the null space, or the nullity of A.

Since every column in A' is either a pivot column or associated with a free variable, and since there are n columns in total, the sum of the number of pivot columns (the rank) and the number of free variables (the nullity) must equal n. Hence,

$$\text{rank}(A) + \text{nullity}(A) = n.$$

This completes the proof of the Rank-Nullity Theorem. □

Example. Recall the previous example, matrix B:

$$B = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 3 & 6 & -3 & 5 \\ 1 & 2 & -1 & -1 \\ 5 & 10 & -5 & 7 \end{pmatrix}$$

The RREF of matrix B is:

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

Since $rank(B) = 2$, by Rank-Nullity theorem, we get $nullity(A) = 4 - rank(A) = 2$. Indeed, we have two free variables, say s, t, and the solutions for $B\mathbf{x} = \mathbf{0}$ are of the form $(t - 2s, s, t, 0)$, and $\{(1, 0, 1, 0), (-2, 1, 0, 0)\}$ is a basis for the null space.

Remark. For an $m \times n$ matrix A with real entries, let us summarize in the table below the essential information relating nullspace(A), rowspace(A), colspace(A), $\mathbb{R}^m$, and $\mathbb{R}^n$:

	Description	Subspace of	Dimension
nullspace(A)	set of vectors x with $Ax = 0$	$\mathbb{R}^n$	nullity(A)
rowspace(A)	span of the row vectors of A	$\mathbb{R}^n$	rank(A)
colspace(A)	span of the column vectors of A	$\mathbb{R}^m$	rank(A)

Notice that rowspace(A) and colspace(A) both have the same dimension, rank(A), but they occur as subspaces of different vectors, namely $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

The following problems are from the textbook (Section 4.9), and they are good for applying rank-nullity theorem.

14. Show that a 3×7 matrix A with nullity$(A) = 4$ must have colspace$(A) = \mathbb{R}^3$. Is rowspace$(A) = \mathbb{R}^3$?
15. Show that a 6×4 matrix A with nullity$(A) = 0$ must have rowspace$(A) = \mathbb{R}^4$. Is colspace$(A) = \mathbb{R}^4$?
16. Prove that if rowspace(A) = nullspace(A), then A contains an even number of columns.

Solution to Problem 14 By the Rank-Nullity Theorem, for a 3×7 matrix A, we have:

$$\text{rank}(A) + \text{nullity}(A) = 7$$

Given that nullity$(A) = 4$, we find:

$$\text{rank}(A) = 7 - 4 = 3$$

Since the rank of A corresponds to the dimension of the column space and A has 3 rows, colspace(A) is at most $\mathbb{R}^3$, and since the rank is 3, it spans $\mathbb{R}^3$. However, even if rowspace(A) is three dimensional, it cannot be $\mathbb{R}^3$ because rowspace(A) is a subspace of $\mathbb{R}^7$.

Solution to Problem 15 For the 6×4 matrix A, if nullity$(A) = 0$, it means there are no free variables, and all columns include leading 1, so:

$$\text{rank}(A) = 4$$

Thus, rowspace(A) is $\mathbb{R}^4$ as it is spanned by 4 linearly independent row vectors. colspace(A) cannot be $\mathbb{R}^4$ because columns are vectors in $\mathbb{R}^6$.

Solution to Problem 16 If rowspace(A) is equal to nullspace(A), then we have rank$(A) =$ nullity(A). By rank-nullity theorem, we have

$$\text{rank}(A) + \text{nullity}(A) = \text{rank}(A) + \text{rank}(A) = 2\text{rank}(A) = \#of columns$$

which makes $\#of columns$ is an even number.