Linear Transformation

Friday 11.08.2023 (Room U3)

1 Linear Transformation

1.1 Definition

A Linear Transformation, T, is a function that maps one vector space V to another vector space W, both over the same field F. It's denoted as T: V → W. A linear transformations must satisfy the following two properties:

1. Additivity: For any vectors u, v in the vector space V, $T(u+v) = T(u) + T(v)$
2. Scalar Multiplication: For any vector v in the vector space V and scalar c in the field F, $T(cv) = cT(v)$

Lemma 1.1 *Linear Transformations Preserve the Origin: Let V and W be vector spaces, and let $T: V \to W$ be a linear transformation. Then T sends the zero vector in V to the zero vector in W. In other words, if $\mathbf{0}_V$ is the zero vector in V, then $T(\mathbf{0}_V) = \mathbf{0}_W$, where $\mathbf{0}_W$ is the zero vector in W.*

Proof 1.2 *Let $\mathbf{0}_V$ be the zero vector in V. By the properties of the zero vector, we have $\mathbf{0}_V = \mathbf{0}_V + \mathbf{0}_V$. Applying the transformation T to both sides of this equation and using the property of additivity, we have $T(\mathbf{0}_V) = T(\mathbf{0}_V + \mathbf{0}_V) = T(\mathbf{0}_V) + T(\mathbf{0}_V)$. Therefore, by subtracting $T(\mathbf{0}_V)$ from both sides, we get $\mathbf{0}_W = T(\mathbf{0}_V)$, where $\mathbf{0}_W$ is the zero vector in W. Hence, $T(\mathbf{0}_V) = \mathbf{0}_W$, as claimed.*

Derivative as a Linear Transformation

Consider the vector spaces $C^0(\mathbb{R})$ and $C^1(\mathbb{R})$, where:

- $C^0(\mathbb{R})$ is the set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$.
- $C^1(\mathbb{R})$ is the set of all functions from $\mathbb{R}$ to $\mathbb{R}$ that have continuous first derivatives.

Now, let $T : C^1(\mathbb{R}) \to C^0(\mathbb{R})$ be the transformation that maps a differentiable function $f(x)$ to its derivative $f'(x)$. The transformation T is linear, as it satisfies the following properties:

1. **Additivity:** For any functions $f(x), g(x) \in C^1(\mathbb{R})$,
$$T(f(x) + g(x)) = f'(x) + g'(x) = T(f(x)) + T(g(x)).$$
2. **Homogeneity:** For any function $f(x) \in C^1(\mathbb{R})$ and any scalar $c \in \mathbb{R}$,
$$T(cf(x)) = cf'(x) = cT(f(x)).$$

Example 1.3 *Let $f(x) = x^2$, $g(x) = \sin x$, and $c = 2$. Then:*

- *Additivity:*
$$\begin{aligned} T(f(x)+g(x)) &= T(x^2+\sin x) = 2x+\cos x \\ T(f(x))+T(g(x)) &= 2x+\cos x, \end{aligned}$$
so $T(f(x)+g(x)) = T(f(x))+T(g(x))$.

- *Homogeneity:*
$$\begin{aligned} T(cf(x)) &= T(2\cdot x^2) = 4x \\ cT(f(x)) &= 2\cdot 2x = 4x, \end{aligned}$$
thus $T(cf(x)) = cT(f(x))$.

These examples demonstrate that the derivative is a linear transformation, satisfying both the additivity and homogeneity properties.

Linearity of the Cross Product Operator

The cross product of two vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$ produces a new vector $\mathbf{c}$ that is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, with a direction given by the right-hand rule. Given a fixed vector $\mathbf{a} \in \mathbb{R}^3$, we can define a linear transformation $T_{\mathbf{a}} : \mathbb{R}^3 \to \mathbb{R}^3$ such that for any $\mathbf{b} \in \mathbb{R}^3$, $T_{\mathbf{a}}(\mathbf{b}) = \mathbf{a} \times \mathbf{b}$.

- **Additivity**:
$$T_{\mathbf{a}}(\mathbf{b}+\mathbf{c}) = (\mathbf{a}\times\mathbf{b}) + (\mathbf{a}\times\mathbf{c}) = T_{\mathbf{a}}(\mathbf{b}) + T_{\mathbf{a}}(\mathbf{c})$$

- **Homogeneity**:
$$T_{\mathbf{a}}(c\cdot\mathbf{b}) = c\cdot(\mathbf{a}\times\mathbf{b}) = c\cdot T_{\mathbf{a}}(\mathbf{b})$$

Example 1.4 *Let* $\mathbf{a} = \begin{bmatrix}1\\2\\3\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}4\\5\\6\end{bmatrix}$, $\mathbf{c} = \begin{bmatrix}7\\8\\9\end{bmatrix}$, *and* $c = 2$. *Then:*

- ***Additivity:***
$$\begin{aligned} T_{\mathbf{a}}(\mathbf{b}+\mathbf{c}) &= \begin{bmatrix}-9\\18\\-9\end{bmatrix} \\ T_{\mathbf{a}}(\mathbf{b}) + T_{\mathbf{a}}(\mathbf{c}) &= \begin{bmatrix}-3\\6\\-3\end{bmatrix} + \begin{bmatrix}-6\\12\\-6\end{bmatrix} = \begin{bmatrix}-9\\18\\-9\end{bmatrix} \end{aligned}$$

- ***Homogeneity:***
$$\begin{aligned} T_{\mathbf{a}}(c\cdot\mathbf{b}) &= \begin{bmatrix}-6\\12\\-6\end{bmatrix} \\ c\cdot T_{\mathbf{a}}(\mathbf{b}) &= \begin{bmatrix}-6\\12\\-6\end{bmatrix} \end{aligned}$$

Example 1.5 *Let T: $\mathbb{R}^2 \to \mathbb{R}^2$ be a **transition** defined by* $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Additivity:

$$T(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}) = T(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}) = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 1 \\ y_1 + y_2 + 2 \end{bmatrix}$$

$$T(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}) + T(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}) = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 2 \\ y_1 + y_2 + 4 \end{bmatrix}$$

T is not a linear transformation because it does not satisfy the additivity property :

$$T(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}) \neq T(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}) + T(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}).$$

2 Matrices and Linear Transformations

The concepts of matrices and linear transformations are two of the most fundamental aspects of linear algebra. The beauty of these two concepts lies in their deep connection: every matrix defines a linear transformation, and conversely, every linear transformation on a finite space can be represented by a matrix.

Example 2.1 *Let T: $\mathbb{R}^2 \to \mathbb{R}^2$ be a **scaling** transformation defined by* $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. *$T$ is a linear transformation because it satisfies both the additivity and scalar multiplication properties.*

Additivity:

$$T(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}) = T(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} = \begin{bmatrix} 2(x_1 + x_2) \\ 3(y_1 + y_2) \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 \\ 3y_1 + 3y_2 \end{bmatrix}$$

$$T(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}) + T(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 3y_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 3y_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 \\ 3y_1 + 3y_2 \end{bmatrix}$$

Scalar Multiplication for any $c \in \mathbb{R}$:

$$T(c\begin{bmatrix} x \\ y \end{bmatrix}) = T(\begin{bmatrix} cx \\ cy \end{bmatrix}) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} cx \\ cy \end{bmatrix} = \begin{bmatrix} 2(cx) \\ 3(cy) \end{bmatrix} = \begin{bmatrix} c(2x) \\ c(3y) \end{bmatrix} = c\begin{bmatrix} 2x \\ 3y \end{bmatrix} = c\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = cT(\begin{bmatrix} x \\ y \end{bmatrix}).$$

Lemma 2.2 *Let A be an $m \times n$ matrix, the $T : \mathbb{R}^n \to \mathbb{R}^m$ that is defined as $T(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$ is a linear transformation.*

Proof 2.3 *We need to show that T satisfies the two properties of a linear transformation:*

1. *Given $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have*

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= A(\mathbf{u} + \mathbf{v}) \\ &= A\mathbf{u} + A\mathbf{v} \quad \textit{(by the distributive property of matrix multiplication)} \\ &= T(\mathbf{u}) + T(\mathbf{v}). \end{aligned}$$

2. Given $\mathbf{u} \in \mathbb{R}^n$ *and a scalar* c*, we have*

$$\begin{aligned} T(c\mathbf{u}) &= A(c\mathbf{u}) \\ &= cA\mathbf{u} \quad \textit{(by the associative property of scalar multiplication)} \\ &= cT(\mathbf{u}). \end{aligned}$$

Therefore, the function $T(\mathbf{x}) = A\mathbf{x}$ *defined by the matrix* A *satisfies both properties of a linear transformation.*

Example 2.4 *Let* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *be a* ***rotation*** *defined by* $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ *is a linear transformation.*

Example 2.5 *Let* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *be a* ***shearing*** *defined by* $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 + kk' & k \\ k' & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$*, where* k *and* k' *are the shear factors, is a linear transformation.*

Example 2.6 *Let* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *be a* ***reflection*** *defined by* $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ *is a linear transformation.*

Example 2.7 *An* ***affine transformation*** *is a linear transformation if and only if the translation vector* $\mathbf{b}$ *is the zero vector. Recall that an affine transformation has the form:*

$$T(\mathbf{v}) = A\mathbf{v} + \mathbf{b}. \tag{1}$$

If $\mathbf{b}$ *equals to the zero vector, then the transformation simplifies to:*

$$T(\mathbf{v}) = A\mathbf{v}, \tag{2}$$

and the affine transformation becomes a linear transformation.

Example 2.8 *Consider the scaling linear transformation* $T : \mathbb{R}^2 \to \mathbb{R}^2$*, defined by the matrix*

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix},$$

where $a \neq b$ *and both* a *and* b *are nonzero.*

This transformation scales space by a factor of a *in the* x*-direction and a factor of* b *in the* y*-direction. If applied to a unit circle, it will stretch or compress the circle into an ellipse.*

The equation for a unit circle centered at the origin is

$$x^2 + y^2 = 1.$$

After applying the transformation T*, we get the equation for an ellipse:*

$$(ax)^2 + (by)^2 = 1.$$

This describes an ellipse centered at the origin, with semi-major and semi-minor axes determined by the values of a *and* b*.*

Proposition 2.9 *Given a linear transformation T from an n-dimensional space to an m-dimensional space ($T : \mathbb{R}^n \to \mathbb{R}^m$), this transformation can be represented as an $m \times n$ matrix.*

Proof 2.10 *Assume $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ denote a basis vectors in $\mathbb{R}^n$.*

Consider the images of these standard basis vectors under the linear transformation T. We have:

$$T(\mathbf{e}_i) = \mathbf{a}_i \quad \text{for } i = 1, 2, \ldots, n,$$

where each $\mathbf{a}_i$ is a vector in $\mathbb{R}^m$.

Let A be the $m \times n$ matrix whose i^{th} column is $\mathbf{a}_i$, that is,

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_n].$$

We claim that A is the matrix representation of T.

To prove this, let $\mathbf{x} \in \mathbb{R}^n$. We can write $\mathbf{x}$ as a linear combination of the basis vectors:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n.$$

Now, consider $T(\mathbf{x})$. Since T is a linear transformation, we can distribute T over the linear combination:

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n) = T(x_1\mathbf{e}_1) + T(x_2\mathbf{e}_2) + \ldots + T(x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \ldots + x_nT(\mathbf{e}_n) = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n. \end{aligned}$$

But the right hand side is exactly the result of multiplying A by $\mathbf{x}$, i.e., $A\mathbf{x}$. Therefore, $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Hence, A is the matrix representation of the linear transformation T.

Example 2.11 *Let's consider a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by the rule $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 3x + 2y \\ x - y \end{bmatrix}$. Here's how we derive the matrix representing T:*

- *Apply T to the standard basis vectors for $\mathbb{R}^2$, which are $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We find $T(e_1) = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $T(e_2) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.*
- *The matrix representing T is then given by placing these transformed vectors as columns in a 2×2 matrix:*
$$[T(e_1) \quad T(e_2)] = \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix}$$

This matrix tells us how to transform any vector from $\mathbb{R}^2$ to another vector in $\mathbb{R}^2$ under the transformation T. We simply perform a matrix-vector multiplication:

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Example 2.12 *Now, let's consider a linear transformation $S : \mathbb{R}^3 \to \mathbb{R}^2$ defined by the rule $S(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x - y \\ y + z \end{bmatrix}$.*

- *Apply S to the standard basis vectors for $\mathbb{R}^3$, which are $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. We find $S(e_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $S(e_2) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and $S(e_3) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.*
- *The matrix representing S is then given by placing these transformed vectors as columns in a 2×3 matrix:*
$$[S(e_1) \quad S(e_2) \quad S(e_3)] = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

This matrix tells us how to transform any vector from $\mathbb{R}^3$ to a vector in $\mathbb{R}^2$ under the transformation S:

$$S(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

2.1 Linear Transformation: Range and Kernel

Consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, where n and m are the dimensions of the input and output vector spaces, respectively. Let A be the $m \times n$ matrix that represents the transformation T.

Range of T:

The range of T, also known as the image of T, is denoted by range(T). It is defined as the set of all possible output vectors that can be obtained by applying T to the input vectors in $\mathbb{R}^n$. The range of T is directly related to the column space of the matrix A, and can be expressed as:

$$\begin{aligned}
\text{range}(T) &= \{T(\mathbf{v}) | \text{ for } \mathbf{v} \in \mathbb{R}^n\} \\
&= \{A\mathbf{v} | \text{ for } \mathbf{v} \in \mathbb{R}^n\} \\
&= \left\{ \begin{bmatrix} a_1, a_2, \cdots, a_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} | \text{ for } \mathbf{v} \in \mathbb{R}^n \right\} \\
&= \{a_1 v_1 + a_2 v_2 + \cdots + a_n v_n | \text{ for } v_1, v_2, \cdots, v_n \in \mathbb{R}\} \\
&= \text{span}(a_1, a_2, \cdots, a_n)
\end{aligned}$$

Here, $a_1, a_2, \ldots, a_n$ are the columns of the $m \times n$ matrix A. Thus, the range is the span of the columns of A, representing all possible linear combinations of these column vectors.

Kernel of T:

The kernel of T, denoted as $\ker(T)$ and also known as the null space of T, consists of all input vectors $\mathbf{v}$ in $\mathbb{R}^n$ that are mapped to the zero vector $\mathbf{0}$ in the output space $\mathbb{R}^m$ under the transformation T. Symbolically, the kernel of T is defined as:

$$\ker(T) = \{\mathbf{v} \in \mathbb{R}^n | T(\mathbf{v}) = \mathbf{0}\} = \{\mathbf{v} \in \mathbb{R}^n | A\mathbf{v} = \mathbf{0}\}$$

The dimension of the kernel, also referred to as the nullity of A, is equal to the number of linearly independent vectors in the kernel.

3 Changing Basis Using Linear Transformations

Changing the basis of a vector or a set of vectors is a pivotal concept in linear algebra. In essence, it entails expressing the same vectors with respect to a new set of basis vectors. If you have a vector in a specific basis and you wish to represent it in another basis, you can utilize a transformation matrix, commonly known as a change of basis matrix.

Let's suppose that $B = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n\}$ represents a basis for a vector space V. Any vector $\mathbf{v}$ in V can be written as a linear combination of the basis vectors. That is to say, $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \ldots + v_n\mathbf{b}_n$, where $v_1, v_2, ..., v_n$ are the coordinates of the vector $\mathbf{v}$ with respect to the basis B.

Now, let's consider another basis $C = \{\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_n\}$ for the same vector space V, and we aim to find the coordinates of $\mathbf{v}$ concerning this new basis. For this, we require a transformation matrix P that correlates the old basis B to the new basis C.

This matrix P has columns representing the coordinates of the new basis vectors (from C) expressed in terms of the old basis B. In essence, the i-th column of P would be the coordinates of $\mathbf{c}_i$ with respect to B.

Formally, we can write:

$$P = \begin{bmatrix} [\mathbf{c}_1]_B & [\mathbf{c}_2]_B & \ldots & [\mathbf{c}_n]_B \end{bmatrix}$$

To represent $\mathbf{v}$ in the new basis C, we multiply the coordinate vector of $\mathbf{v}$ in the old basis B (that is, $[v]_B = [v_1, v_2, ..., v_n]^T$) by the inverse of the transformation matrix P.

$$[v]_C = P^{-1}[v]_B$$

This resultant vector $[v]_C$ has the same "meaning" as $\mathbf{v}$, but is expressed concerning the new basis C.

Example 3.1 *Assume we have two bases, B and C, for $\mathbb{R}^2$. These bases are defined as follows:*

$$B = \left\{ b_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, b_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \quad \textit{and} \quad C = \left\{ c_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, c_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Our goal is to represent the vector $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ *in both bases* B *and* C.

In basis B*, the vector is represented as follows:*

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Rightarrow [v]_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Now, to express the same vector in basis C*, we write:*

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow [v]_C = \begin{bmatrix} x \\ y \end{bmatrix}$$

In order to find values for x *and* y*, we first construct a matrix* P *that has as its columns the coordinates of the basis vectors from* C *in terms of basis* B*. Since* C *is already expressed in terms of* B*, matrix* P *is given by:*

$$\begin{cases} c_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = b_1 - 2b_2 \\ c_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = b_1 \end{cases} \Rightarrow P = \begin{bmatrix} 1 & 1 \\ -2 & 0 \end{bmatrix}$$

Next, we calculate the inverse of matrix P*, denoted* P^{-1}*:*

$$P^{-1} = \frac{1}{2} \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}$$

The final step is to find the coordinates of $\mathbf{v}$ *in the new basis* C*:*

$$[v]_C = P^{-1}[v]_B = \frac{1}{2} \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 2.5 \end{bmatrix}$$

This computation gives us:

$$-0.5c_1 + 2.5c_2 = -0.5 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 2.5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Therefore, the coordinates of vector $\mathbf{v}$ *in the new basis* C *are* -0.5 *and* 2.5*. Thus, if we have the coefficients of a vector* $\mathbf{v}$ *when it is expressed in basis* B *(denoted as* $[v]_B$*), we can find its coefficients when expressed in basis* C *(denoted as* $[v]_C$*)with the following relation:*

$$[v]_C = P^{-1}[v]_B$$

Example 3.2 *Let's consider two distinct bases,* F *and* H*, for* $\mathbb{R}^4$*. These are defined as follows:*

$$F = \left\{ f_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, f_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, f_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, f_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\}$$

$$H = \left\{ h_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, h_2 = \begin{bmatrix}1\\-1\\0\\0\end{bmatrix}, h_3 = \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, h_4 = \begin{bmatrix}0\\0\\1\\-1\end{bmatrix} \right\}$$

Suppose we aim to change the basis from F to H. We first express each h_i as a combination of f_i:

$$\begin{cases} h_1 = f_1 + f_3 \\ h_2 = f_2 \\ h_3 = f_1 - f_3 \\ h_4 = f_4 \end{cases} \Rightarrow P = \begin{bmatrix}1 & 0 & 1 & 0\\0 & 1 & 0 & 0\\1 & 0 & -1 & 0\\0 & 0 & 0 & 1\end{bmatrix}$$

Next, we calculate P^{-1}:

$$P^{-1} = \frac{1}{2}\begin{bmatrix}1 & 0 & 1 & 0\\0 & 2 & 0 & 0\\1 & 0 & -1 & 0\\0 & 0 & 0 & 2\end{bmatrix}$$

Let's confirm our calculations by considering a vector $\mathbf{v}$ in $\mathbb{R}^4$:

$$\mathbf{v} = \begin{bmatrix}2\\3\\5\\7\end{bmatrix}$$

We can express $\mathbf{v}$ in terms of the basis F as follows:

$$2.5\begin{bmatrix}1\\1\\0\\0\end{bmatrix} - 0.5\begin{bmatrix}1\\-1\\0\\0\end{bmatrix} + 6\begin{bmatrix}0\\0\\1\\1\end{bmatrix} - 1\begin{bmatrix}0\\0\\1\\-1\end{bmatrix} = \begin{bmatrix}2\\3\\5\\7\end{bmatrix} \quad \Rightarrow \quad 2.5f_1 - 0.5f_2 + 6f_3 - f_4 = \mathbf{v} \quad \Rightarrow \quad [v]_F = \begin{bmatrix}2.5\\-0.5\\6\\-1\end{bmatrix}$$

Now we want to express $\mathbf{v}$ in terms of the basis H. We need to find coefficients a, b, c, and d such that :

$$\mathbf{v} = ah_1 + bh_2 + ch_3 + dh_4$$

By applying the change of basis formula, we find the coefficients a, b, c, and d:

$$[v]_H = P^{-1}[v]_F = \frac{1}{2}\begin{bmatrix}1 & 0 & 1 & 0\\0 & 2 & 0 & 0\\1 & 0 & -1 & 0\\0 & 0 & 0 & 2\end{bmatrix}\begin{bmatrix}2.5\\-0.5\\6\\-1\end{bmatrix} = \begin{bmatrix}4.25\\-0.5\\-1.75\\-1\end{bmatrix}$$

Therefore, the coordinates of the vector $\mathbf{v}$ *in the basis* H *are* 4.25, -0.5, -1.75, *and* -1. *We can confirm this by expressing* $\mathbf{v}$ *in terms of* H*:*

$$4.25\begin{bmatrix}1\\1\\1\\1\end{bmatrix} - 0.5\begin{bmatrix}1\\-1\\0\\0\end{bmatrix} - 1.75\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} - 1\begin{bmatrix}0\\0\\1\\-1\end{bmatrix} = \begin{bmatrix}2\\3\\5\\7\end{bmatrix} \quad \Rightarrow \quad 4.25h_1 - 0.5h_2 - 1.75h_3 - h_4 = \mathbf{v},$$

demonstrating the successful change of basis from F *to* H*.*

Example 3.3 *Let's consider two distinct bases, the standard basis* E *and a non-standard basis* B*, for* $\mathbb{R}^3$*. These are defined as follows:*

$$E = \left\{ e_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, e_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, e_3 = \begin{bmatrix}0\\0\\1\end{bmatrix} \right\}.$$

$$B = \left\{ b_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, b_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}, b_3 = \begin{bmatrix}1\\0\\1\end{bmatrix} \right\}$$

To convert from the standard basis E *to the non-standard basis* B*, we initially express each basis vector in* B *in terms of the basis vectors in* E*. This gives us:*

$$\begin{aligned} b_1 &= e_1 + e_2, \\ b_2 &= e_2 + e_3, \\ b_3 &= e_1 + e_3. \end{aligned}$$

From these relationships, we obtain the change-of-basis matrix P *from* E *to* B *as:*

$$P = \begin{bmatrix}1 & 0 & 1\\1 & 1 & 0\\0 & 1 & 1\end{bmatrix}.$$

The inverse of this matrix, P^{-1}*, represents the transition from the basis* B *back to the basis* E*:*

$$P^{-1} = \frac{1}{2}\begin{bmatrix}1 & 1 & -1\\-1 & 1 & 1\\1 & -1 & 1\end{bmatrix}.$$

Now, suppose we have a vector $\mathbf{v}$ *in* $\mathbb{R}^3$*, represented as:*

$$\mathbf{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}$$

and it is defined in the standard basis E*. To convert this vector to the non-standard basis* B*, we apply the inverse matrix* P^{-1}*:*

$$[v]_B = P^{-1}[v]_E = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

Therefore, the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$*, expressed in the basis* B*, is* $\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$.

We can validate our results by representing $\mathbf{v}$ *in terms of the vectors in* B*:*

$$\mathbf{v} = 0b_1 + 2b_2 + b_3 = 0\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

which confirms the successful change of the basis for the vector.