Matrix Representations of Linear Transformations and Changes of Coordinates

0.1 Subspaces and Bases

0.1.1 Definitions

A subspace \( V \) of \( \mathbb{R}^n \) is a subset of \( \mathbb{R}^n \) that contains the zero element and is closed under addition and scalar multiplication:

1. \( 0 \in V \)
2. \( u, v \in V \implies u + v \in V \)
3. \( u \in V \) and \( k \in \mathbb{R} \implies ku \in V \)

Equivalently, \( V \) is a subspace if \( au + bv \in V \) for all \( a, b \in \mathbb{R} \) and \( u, v \in V \). (You should try to prove that this is an equivalent statement to the first.)

Example 0.1 Let \( V = \{(t, 3t, -2t) \mid t \in \mathbb{R}\} \). Then \( V \) is a subspace of \( \mathbb{R}^3 \):

1. \( 0 \in V \) because we can take \( t = 0 \).
2. If \( u, v \in V \), then \( u = (s, 3s, -2s) \) and \( v = (t, 3t, -2t) \) for some real numbers \( s \) and \( t \). But then \( u + v = (s + t, 3s + 3t, -2s - 2t) = (s + t, 3(s + t), -2(s + t)) = (t', 3t', -2t') \in V \)
   where \( t' = s + t \in \mathbb{R} \).
3. If \( u \in V \), then \( u = (t, 3t, -2t) \) for some \( t \in \mathbb{R} \), so if \( k \in \mathbb{R} \), then \( ku = (kt, 3kt, -2kt) = (t', 3t', -2t') \in V \)
   where \( t' = kt \in \mathbb{R} \). ■

Example 0.2 The unit circle \( S^1 \) in \( \mathbb{R}^2 \) is not a subspace because it doesn’t contain \( 0 = (0,0) \) and because, for example, \( (1,0) \) and \( (0,1) \) lie in \( S \) but \( (1,0) + (0,1) = (1,1) \) does not. Similarly, \( (1,0) \) lies in \( S \) but \( 2(1,0) = (2,0) \) does not. ■

A linear combination of vectors \( v_1, \ldots, v_k \in \mathbb{R}^n \) is the finite sum

\[
a_1 v_1 + \cdots + a_k v_k
\]

which is a vector in \( \mathbb{R}^n \) (because \( \mathbb{R}^n \) is a subspace of itself, right?). The \( a_i \in \mathbb{R} \) are called the coefficients of the linear combination. If \( a_1 = \cdots = a_k = 0 \), then the linear combination is said to be trivial. In particular, considering the special case of \( 0 \in \mathbb{R}^n \), the zero vector, we note that \( 0 \) may always be represented as a linear combination of any vectors \( u_1, \ldots, u_k \in \mathbb{R}^n \),

\[
0u_1 + \cdots + 0u_k = 0
\]

This representation is called the trivial representation of \( 0 \) by \( u_1, \ldots, u_k \). If, on the other hand, there are vectors \( u_1, \ldots, u_k \in \mathbb{R}^n \) and scalars \( a_1, \ldots, a_n \in \mathbb{R} \) such that

\[
a_1 u_1 + \cdots + a_k u_k = 0
\]
where at least one \( a_i \neq 0 \), then that linear combination is called a \textbf{nontrivial representation of} \( 0 \). Using linear combinations we can generate subspaces, as follows. If \( S \) is a nonempty subset of \( \mathbb{R}^n \), then the \textbf{span} of \( S \) is given by

\[
\text{span}(S) := \{ v \in \mathbb{R}^n \mid v \text{ is a linear combination of vectors in } S \} \quad (0.2)
\]

The span of the empty set, \( \emptyset \), is by definition

\[
\text{span}(\emptyset) := \{ 0 \} \quad (0.3)
\]

\textbf{Remark 0.3} We showed in class that \( \text{span}(S) \) is always a subspace of \( \mathbb{R}^n \) (well, we showed this for \( S \) a finite collection of vectors \( S = \{ u_1, \ldots, u_k \} \), but you should check that it’s true for any \( S \)). ■

Let \( V := \text{span}(S) \) be the subspace of \( \mathbb{R}^n \) spanned by some \( S \subseteq \mathbb{R}^n \). Then \( S \) is said to generate or \( \text{span} \) \( V \), and to be a \textbf{generating} or \textbf{spanning set} for \( V \). If \( V \) is already known to be a subspace, then finding a spanning set \( S \) for \( V \) can be useful, because it is often easier to work with the smaller spanning set than with the entire subspace \( V \), for example if we are trying to understand the behavior of linear transformations on \( V \).

\textbf{Example 0.4} Let \( S \) be the unit circle in \( \mathbb{R}^3 \) which lies in the \( x-y \) plane. Then \( \text{span}(S) \) is the entire \( x-y \) plane. ■

\textbf{Example 0.5} Let \( S = \{(x, y, z) \in \mathbb{R}^3 \mid x = y = 0, \ 1 < z < 3 \} \). Then \( \text{span}(S) \) is the \( z \)-axis. ■

A nonempty subset \( S \) of a vector space \( \mathbb{R}^n \) is said to be \textbf{linearly independent} if, taking any finite number of distinct vectors \( u_1, \ldots, u_k \in S \), we have for all \( a_1, \ldots, a_k \in \mathbb{R} \) that

\[
a_1 u_1 + a_2 u_2 + \cdots + a_k u_k = 0 \quad \Rightarrow \quad a_1 = \cdots = a_k = 0
\]

That is \( S \) is linearly independent if the \textit{only} representation of \( 0 \in \mathbb{R}^n \) by vectors in \( S \) is the trivial one. In this case, the vectors \( u_1, \ldots, u_k \) themselves are also said to be linearly independent. Otherwise, if there is at least one nontrivial representation of \( 0 \) by vectors in \( S \), then \( S \) is said to be \textbf{linearly dependent}.

\textbf{Example 0.6} The vectors \( u = (1, 2) \) and \( v = (0, -1) \) in \( \mathbb{R}^2 \) are linearly independent, because if

\[
a u + b v = 0
\]

that is

\[
a(1, 2) + b(0, -1) = (0, 0)
\]

then \( (a, 2a - b) = (0, 0) \), which gives a system of equations:

\[
\begin{align*}
a & = 0 \\
2a - b & = 0
\end{align*}
\]

or

\[
\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

But the matrix \( \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \) is invertible, in fact it is its own inverse, so that left-multiplying both sides by it gives

\[
\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}^2 \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

which means \( a = b = 0 \). ■
Example 0.7  The vectors $(1, 2, 3), (4, 5, 6), (7, 8, 9) \in \mathbb{R}^3$ are not linearly independent because

\[ 1(1, 2, 3) - 2(4, 5, 6) + 1(7, 8, 9) = (0, 0, 0) \]

That is, we have found $a = 1$, $b = -2$ and $c = 1$, not all of which are zero, such that $a(1, 2, 3) + b(4, 5, 6) + c(7, 8, 9) = (0, 0, 0)$. □

Given $\emptyset \neq S \subseteq V$, a nonzero vector $v \in S$ is said to be an **essentially unique linear combination** of the vectors in $S$ if, up to order of terms, there is one and only one way to express $v$ as a linear combination of $u_1, \ldots, u_k \in S$. That is, if there are $a_1, \ldots, a_n, b_1, \ldots, b_\ell \in \mathbb{R}\setminus\{0\}$ and distinct $u_1, \ldots, u_k \in S$ and distinct $v_1, \ldots, v_\ell \in S$ distinct, then, re-indexing the $b_i$s if necessary,

\[
\begin{align*}
v &= a_1u_1 + \cdots + a_nu_k \\
&= b_1v_1 + \cdots + b_\ell v_\ell
\end{align*}
\]

\[= k = \ell \quad \text{and} \quad \begin{cases} a_i = b_i \\ u_i = v_i \end{cases} \quad \text{for all } i = 1, \ldots, k
\]

If $V$ is a subspace of $\mathbb{R}^n$, then a subset $\beta$ of $V$ is called a **basis** for $V$ if it is linearly independent and spans $V$. We also say that the **vectors** of $\beta$ form a basis for $V$. Equivalently, as explained in Theorem 0.11 below, $\beta$ is a basis if every nonzero vector $v \in V$ is an essentially unique linear combination of vectors in $\beta$.

Remark 0.8  In the context of inner product spaces $V$ of infinite dimension, there is a difference between a vector space basis, the **Hamel basis** of $V$, and an orthonormal basis for $V$, the **Hilbert basis** for $V$, because though the two always exist, they are not always equal unless $\dim(V) < \infty$. □

The **dimension** of a subspace $V$ of $\mathbb{R}^n$ is the cardinality of any basis for $V$, i.e. the number of elements in $\beta$ (which may in principle be infinite), and is denoted $\dim(V)$. This is a well defined concept, by Theorem 0.13 below, since all bases have the same size. $V$ is **finite-dimensional** if it is the zero vector space $\{0\}$ or if it has a basis of finite cardinality. Otherwise, if it’s basis has infinite cardinality, it is called **infinite-dimensional**. In the former case, $\dim(V) = |\beta| = k < \infty$ for some $n \in \mathbb{N}$, and $V$ is said to be $k$-**dimensional**, while in the latter case, $\dim(V) = |\beta| = \kappa$, where $\kappa$ is a cardinal number, and $V$ is said to be $\kappa$-**dimensional**.

Remark 0.9  Bases are not unique. For example, $\beta = \{e_1, e_2\}$ and $\gamma = \{(1, 1), (1, 0)\}$ are both bases for $\mathbb{R}^2$. □

If $V$ is finite-dimensional, say of dimension $n$, then an **ordered basis** for $V$ a finite sequence or $n$-tuple $(v_1, \ldots, v_n)$ of linearly independent vectors $v_1, \ldots, v_n \in V$ such that $\{v_1, \ldots, v_n\}$ is a basis for $V$. If $V$ is infinite-dimensional but with a countable basis, then an ordered basis is a sequence $(v_n)_{n \in \mathbb{N}}$ such that the set $\{v_n \mid n \in \mathbb{N}\}$ is a basis for $V$. 
0.1.2 Properties of Bases

Theorem 0.10 Vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n \) are linearly independent iff no \( \mathbf{v}_i \) is a linear combination of the other \( \mathbf{v}_j \).

Proof: Let \( \mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n \) be linearly independent and suppose that \( \mathbf{v}_k = c_1 \mathbf{v}_1 + \cdots + c_{k-1} \mathbf{v}_{k-1} \) (we may suppose \( \mathbf{v}_k \) is a linear combination of the other \( \mathbf{v}_j \), else we can simply re-index so that this is the case). Then
\[
c_1 \mathbf{v}_1 + \cdots + c_{k-1} \mathbf{v}_{k-1} + (-1) \mathbf{v}_k = 0
\]
But this contradicts linear independence, since \( -1 \neq 0 \). Hence \( \mathbf{v}_k \) cannot be a linear combination of the other \( \mathbf{v}_j \). By re-indexing the \( \mathbf{v}_i \) we can conclude this for all \( \mathbf{v}_i \).

Conversely, suppose \( \mathbf{v}_1, \ldots, \mathbf{v}_k \) are linearly dependent, i.e. there are scalars \( c_1, \ldots, c_k \in \mathbb{R} \) not all zero such that
\[
c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k = 0
\]
Say \( c_i \neq 0 \). Then,
\[
\mathbf{v}_i = \left( -\frac{c_1}{c_i} \right) \mathbf{v}_1 + \cdots + \left( -\frac{c_{i-1}}{c_i} \right) \mathbf{v}_{i-1} + \left( -\frac{c_{i+1}}{c_i} \right) \mathbf{v}_{i+1} + \cdots + \left( -\frac{c_k}{c_i} \right) \mathbf{v}_k
\]
so that \( \mathbf{v}_i \) is a linear combination of the other \( \mathbf{v}_j \). This is the contrapositive of the equivalent statement, "If no \( \mathbf{v}_i \) is a linear combination of the other \( \mathbf{v}_j \), then \( \mathbf{v}_1, \ldots, \mathbf{v}_k \) are linearly independent.” ■

Theorem 0.11 Let \( V \) be a subspace of \( \mathbb{R}^n \). Then a collection \( \beta = \{ \mathbf{v}_1, \ldots, \mathbf{v}_k \} \) is a basis for \( V \) iff every vector \( \mathbf{v} \in V \) has an essentially unique expression as a linear combination of the basis vectors \( \mathbf{v}_i \).

Proof: Suppose \( \beta \) is a basis and suppose that \( \mathbf{v} \) has two representations as a linear combination of the \( \mathbf{v}_i \):
\[
\mathbf{v} = c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k = d_1 \mathbf{v}_1 + \cdots + d_k \mathbf{v}_k
\]
Then,
\[
0 = \mathbf{v} - \mathbf{v} = (c_1 - d_1) \mathbf{v}_1 + \cdots + (c_k - d_k) \mathbf{v}_k
\]
so by linear independence we must have \( c_1 - d_1 = \cdots = c_k - d_k = 0 \), or \( c_i = d_i \) for all \( i \), and so \( \mathbf{v} \) has only one expression as a linear combination of basis vectors, up to order of the \( \mathbf{v}_i \).

Conversely, suppose every \( \mathbf{v} \in V \) has an essentially unique expression as a linear combination of the \( \mathbf{v}_i \). Then clearly \( \beta \) is a spanning set for \( V \), and moreover the \( \mathbf{v}_i \) are linearly independent: for note, since \( 0 \mathbf{v}_1 + \cdots + 0 \mathbf{v}_k = \mathbf{0} \), by uniqueness of representations we must have \( c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k = \mathbf{0} \implies c_1 = \cdots = c_k = 0 \). Thus \( \beta \) is a basis. ■
Theorem 0.12 (Replacement Theorem) Let $V$ be a subspace of $\mathbb{R}^n$ and let $v_1, \ldots, v_p$ and $w_1, \ldots, w_q$ be vectors in $V$. If $v_1, \ldots, v_p$ are linearly independent and $w_1, \ldots, w_q$ span $V$, then $p \leq q$.

Proof: Let $A = [w_1 \cdots w_q] \in \mathbb{R}^{n \times q}$ be the matrix whose columns are the $w_j$ and let $B = [v_1 \cdots v_p] \in \mathbb{R}^{n \times p}$ be the matrix whose columns are the $v_k$. Then note that

$$\{v_1, \ldots, v_p\} \subseteq V = \text{span}(w_1, \ldots, w_q) = \text{im} A$$

Thus, there exist $u_1, \ldots, u_p \in \mathbb{R}^q$ such that $Au_i = v_i$. Consequently,

$$B = [v_1 \cdots v_p] = [Au_1 \cdots u_p] = A[u_1 \cdots u_p] = AC$$

where $C = [u_1 \cdots u_p] \in \mathbb{R}^{q \times p}$. Now, since $v_1 \cdots v_p$ are linearly independent, $c_1v_1 + \cdots + c_pv_p = 0$ implies all $c_i = 0$, i.e. $Ba = 0$ implies $c = 0$, or $\ker B = \{0\}$. But you will notice that $\ker C \subseteq \ker B$, since if $x \in \ker C$, then the fact that $B = AC$ implies $Bx = (AC)x = A(Cx) = A0 = 0$, or $x \in \ker C$. Since $\ker B = \{0\}$, this means that $\ker C = \{0\}$ as well. But then $C$ must have at least as many rows as columns, i.e. $p \leq q$, because $rref(C)$ must have the form $[I_p \ O]$, possibly with no $O$ submatrix, but at least with $I_p$ in the top portion. \hfill \blacksquare

Theorem 0.13 Let $V$ be a subspace for $\mathbb{R}^n$. Then all bases for $V$ have the same size.

Proof: By the previous theorem two bases $\beta = \{v_1, \ldots, v_p\}$ and $\gamma = \{w_1, \ldots, w_q\}$ for $V$ both span $V$ and both are linearly independent, so we have $p \leq q$ and $p \geq q$. Therefore $p = q$. \hfill \blacksquare

Corollary 0.14 All bases for $\mathbb{R}^n$ have $n$ vectors.

Proof: Notice that $\rho = \{e_1, \ldots, e_n\}$ forms a basis for $\mathbb{R}^n$: first, the elementary vectors $e_i$ span $\mathbb{R}^n$, since if $x = (a_1, \ldots, a_n) \in \mathbb{R}^n$, then

$$x = (a_1, \ldots, a_n) = a_1(1,0,\ldots,0) + a_2(0,1,0,\ldots,0) + \cdots + a_n(0,\ldots,0,1)$$

$$= a_1e_1 + a_2e_2 + \cdots + a_ne_n \in \text{span}(e_1, \ldots, e_n)$$

Also, $e_1, \ldots, e_n$ are linearly independent, for if

$$0 = c_1e_1 + \cdots + c_ne_n = \begin{bmatrix} c_1 \\
\vdots \\
c_n \end{bmatrix} = I_n c$$

then $c = (c_1, \ldots, c_n) \in \ker I_n = \{0\}$, so $c_1 = \cdots = c_n = 0$. Since $|\rho| = n$, all bases $\beta$ for $\mathbb{R}^n$ satisfy $|\beta| = n$ be the previous theorem. \hfill \blacksquare

Theorem 0.15 (Characterizations of Bases) If $V$ is a subspace of $\mathbb{R}^n$ and $\dim(V) = k$, then

1. There are at most $k$ linearly independent vectors in $V$. Consequently, a basis is a maximal linearly independent set in $V$.
2. At least $k$ vectors are needed to span $V$. Thus a basis is a minimal spanning set.
3. If $k$ vectors in $V$ are linearly independent, then they form a basis for $V$.
4. If $k$ vectors span $V$, then they form a basis for $V$. 


**Proof:** (1) If \( v_1, \ldots, v_p \in V \) are linearly independent and \( w_1, \ldots, w_k \in V \) form a basis for \( V \), then \( p \leq k \) by the Replacement Theorem. (2) If \( v_1, \ldots, v_p \in V \) span \( V \) and \( w_1, \ldots, w_k \in V \) form a basis for \( V \), then again we must have \( k \leq p \) by the Replacement Theorem. (3) If \( v_1, \ldots, v_k \in V \) are linearly independent, we must show they also span \( V \). Pick \( v \in V \) and note that by (1) the vectors \( v_1, \ldots, v_k, \subseteq \in V \) are linearly dependent, because there are \( k + 1 \) of them. (4) If \( v_1, \ldots, v_k, \subseteq \in V \) span \( V \) but are not linearly independent, then say \( v_k \in \text{span}(v_1, \ldots, v_{k-1}) \). But in this case \( V = \text{span}(v_1, \ldots, v_{k-1}) \), contradicting (2).

**Theorem 0.16**  If \( A \in \mathbb{R}^{m \times n} \), then \( \dim(\text{im} \, A) = \text{rank} \, A \).

**Proof:** This follows from Theorem 0.15 in *Systems of Linear Equations*, since if \( B = \text{rref}(A) \), then \( \text{rank} \, A = \text{rank} \, B = \# \text{ of columns of the form } e_i \text{ in } B \) = \# of nonredundant vectors in \( A \).

**Theorem 0.17 (Rank-Nullity Theorem)**  If \( A \in \mathbb{R}^{m \times n} \), then

\[
\dim(\ker A) + \dim(\text{im} \, A) = n \tag{0.4}
\]

or

\[
\text{null} \, A + \text{rank} \, A = n \tag{0.5}
\]

**Proof:** If \( B = \text{rref}(A) \), then \( \dim(\ker A) = n - \# \text{ of leading } 1s \) = \( n - \text{rank} \, A \).
0.2 Coordinate Representations of Vectors and Matrix Representations of Linear Transformations

0.2.1 Definitions

If \( \beta = (v_1, \ldots, v_k) \) is an ordered basis for a subspace \( V \) of \( \mathbb{R}^n \), then we know that for any vector \( v \in V \) there are unique scalars \( a_1, \ldots, a_k \in \mathbb{R} \) such that
\[
v = a_1 v_1 + \cdots + a_k v_k
\]
The coordinate vector of \( v \in \mathbb{R}^n \) with respect to, or relative to, \( \beta \) is defined to be the (column) vector in \( \mathbb{R}^k \) consisting of the scalars \( a_i \):
\[
[x]_\beta := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix} 
\]
and the coordinate map, also called the standard representation of \( V \) with respect to \( \beta \),
\[
\phi_\beta : V \to \mathbb{R}^k
\]
is given by
\[
\phi_\beta(x) = [x]_\beta
\]

Example 0.18 Let \( v = (5, 7, 9) \in \mathbb{R}^3 \) and let \( \beta = (v_1, v_2) \) be the ordered basis for \( V = \text{span}(v_1, v_2) \), where \( v_1 = (1, 1, 1) \) and \( v_2 = (1, 2, 3) \). Can you express \( v \) as a linear combination of \( v_1 \) and \( v_2 \)? In other words, does \( v \) lie in \( V \)? If so, find \([v]_\beta\).

Solution: To find out whether \( v \) lies in \( V \), we must see if there are scalars \( a, b \in \mathbb{R} \) such that
\[
v = av_1 + bv_2 = [v_1 \ v_2] \begin{pmatrix} a \\ b \end{pmatrix}
\]
or
\[
\begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}
\]
This is a system of equations, \( Ax = b \) with
\[
A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \quad x = \begin{pmatrix} a \\ b \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix}
\]
Well, the augmented matrix \([A|b]\) reduces to \( \text{rref}([A|b]) \) as follows:
\[
\begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 7 \\ 1 & 3 & 9 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}
\]
This means that \( a = 3 \) and \( b = 2 \), so that
\[
v = 3v_1 + 2v_2
\]
and \( v \) lies in \( V \), and moreover
\[
[v]_\beta = \begin{pmatrix} 3 \\ 2 \end{pmatrix}
\]
In general, if $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is a basis for a subspace $V$ of $\mathbb{R}^n$ and $\mathbf{v} \in V$, then the coordinate map will give us a matrix equation if we treat all the vectors $\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_k$ as column vectors:

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k = [\mathbf{v}_1 \ldots \mathbf{v}_k] \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$$

or

$$\mathbf{v} = B[\mathbf{v}]_\beta$$

where $B = [\mathbf{v}_1 \ldots \mathbf{v}_k]$, the $n \times k$ matrix with columns $\mathbf{v}_j$. If $V = \mathbb{R}^n$, then $B$ will be an $n \times n$ matrix whose columns are linearly independent. Therefore, $\text{im} B = \mathbb{R}^n$, so that by the Rank-Nullity Theorem $\ker B = \{\mathbf{0}\}$, which means $B$ represents an injective and surjective linear transformation, and is therefore invertible. In this case, we can solve for $[\mathbf{v}]_\beta$ rather easily:

$$[\mathbf{v}]_\beta = B^{-1} \mathbf{v} \quad (0.9)$$

Let $V$ and $W$ be finite dimensional subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, with ordered bases $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_k)$ and $\gamma = (\mathbf{w}_1, \ldots, \mathbf{w}_\ell)$, respectively. If there exist (and there do exist) unique scalars $a_{ij} \in \mathbb{R}$ such that

$$T(\mathbf{v}_j) = \sum_{i=1}^\ell a_{ij} \mathbf{w}_i \quad \text{for } j = 1, \ldots, k \quad (0.10)$$

then the **matrix representation of a linear transformation** $T \in \mathcal{L}(V,W)$ in the ordered bases $\beta$ and $\gamma$ is the $\ell \times k$ matrix $A$ defined by $A_{ij} = a_{ij}$,

$$A = [T]_\beta^\gamma := [T(\mathbf{v}_1)]_\gamma [T(\mathbf{v}_2)]_\gamma \cdots [T(\mathbf{v}_k)]_\gamma$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\ell 1} & a_{\ell 2} & \cdots & a_{\ell k} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{\ell 1} & \cdots & a_{\ell k} \end{bmatrix}$$

Note that $[T(\mathbf{v})]_\gamma = \varphi_\gamma(T(\mathbf{v})) = (\varphi_\gamma \circ T)(\mathbf{v})$.

**Notation 0.19** If $V = W$ and $\beta = \gamma$, we write $[T]_\beta$ instead of $[T]_\beta^\beta$. \hfill \blacksquare

**Example 0.20** Let $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ be given by $T(x,y) = (x+y, 2x-y, 3x+5y)$. In terms of matrices and column vectors $T$ behaves as follows:

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

But this matrix, call it $A$, is actually the representation of $T$ with respect to the standard ordered bases $\rho_2 = (\mathbf{e}_1, \mathbf{e}_2)$ and $\rho_3 = (\mathbf{e}_1, \mathbf{e}_2, 3)$, that $A = [T]_{\rho_2}^{\rho_3}$. What if we were to choose different bases for $\mathbb{R}^2$ and $\mathbb{R}^3$? Say,

$$\beta = ((1,1), (0,-1)), \quad \gamma = ((1,1,1), (1,0,1), (0,0,1))$$

$$\rho_2 = (\mathbf{e}_1, \mathbf{e}_2), \quad \rho_3 = (\mathbf{e}_1, \mathbf{e}_2, 3)$$

But this matrix, call it $A$, is actually the representation of $T$ with respect to the standard ordered bases $\rho_2 = (\mathbf{e}_1, \mathbf{e}_2)$ and $\rho_3 = (\mathbf{e}_1, \mathbf{e}_2, 3)$, that $A = [T]_{\rho_2}^{\rho_3}$. What if we were to choose different bases for $\mathbb{R}^2$ and $\mathbb{R}^3$? Say,
How would $T$ look with respect to these bases? Let us first find the coordinate representations of $T(1,1)$ and $T(0,−1)$ with respect to $γ$: Note, $T(1,1) = (2,1,8)$ and $T(0,−1) = (−1,1,−5)$, and to find $[T(1,1)]_γ$ and $[T(0,−1)]_γ$ we have to solve the equations:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}_γ \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}_γ \begin{bmatrix} −1 \\ 1 \\ −5 \end{bmatrix} = \begin{bmatrix} −1 \\ 1 \\ −5 \end{bmatrix}$$

If $B$ is the matrix

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

then $B^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & −1 & 0 \\ −1 & 0 & 1 \end{bmatrix}$, so

$$\begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix}_γ = \begin{bmatrix} 0 & 1 & 0 \\ 1 & −1 & 0 \\ −1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 6 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} −1 \\ 1 \\ −5 \end{bmatrix}_γ = \begin{bmatrix} 0 & 1 & 0 \\ 1 & −1 & 0 \\ −1 & 0 & 1 \end{bmatrix} \begin{bmatrix} −1 \\ 1 \\ −5 \end{bmatrix} = \begin{bmatrix} −2 \\ −4 \end{bmatrix}$$

Let us verify this:

$$1(1,1,1) + 1(0,1,0) + 6(0,0,1) = (2,1,8) \quad \text{and} \quad 1(1,1,1) − 2(1,0,1) − 4(0,0,1) = (−1,1,−5)$$

so indeed we have found $[T(1,1)]_γ$ and $[T(0,−1)]_γ$, and therefore

$$[T]_β := \begin{bmatrix} [T(1,1)]_γ & [T(0,−1)]_γ \end{bmatrix} = \begin{bmatrix} 1 & −2 \\ 1 & 1 \\ 1 & −2 \\ 6 & −4 \end{bmatrix}$$

0.2.2 Properties of Coordinate and Matrix Representations

**Theorem 0.21 (Linearity of Coordinates)** Let $β$ be a basis for a subspace $V$ of $\mathbb{R}^n$. Then, for all $x, y ∈ V$ and all $k ∈ \mathbb{R}$ we have

1. $[x + y]_β = [x]_β + [y]_β$
2. $[kx]_β = k[x]_β$

**Proof:** (1) On the one hand, $x + y = B[x + y]_β$, and on the other $x = B[x]_β$ and $y = B[y]_β$, so $x + y = B[x]_β + B[y]_β$. Thus,

$$B[x + y]_β = x + y = B[x]_β + B[y]_β = B([x]_β + [y]_β)$$

so that, subtracting the right hand side from both sides, we get

$$B([x + y]_β − ([x]_β + [y]_β)) = 0$$

Now, $B$’s columns are basis vectors, so they are linearly independent, which means $Bx = 0$ has only the trivial solution, because if $β = (v_1, \ldots, v_k)$ and $x = (x_1, \ldots, x_k)$, then $0 = x_1v_1 + \cdots + x_kv_k = Bx \implies x_1 = \cdots = x_k = 0$, or $x = 0$. But this means the kernel of $B$ is $\{0\}$, so that

$$[x + y]_β − ([x]_β + [y]_β) = 0$$

or

$$[x + y]_β = [x]_β + [y]_β$$

The proof of (2) follows even more straightforwardly: First, note that if $[x]_β = [a_1 \cdots a_k]^T$, then $x = a_1v_1 + \cdots + a_kv_k$, so that $kx = ka_1v_1 + \cdots + ka_kv_k$, and therefore

$$[kx]_β = [ka_1 \cdots ka_n]^T = k[a_1 \cdots a_k]^T = k[x]_β$$
Corollary 0.22 \ The coordinate maps $\varphi_\beta$ are linear, i.e. $\varphi_\beta \in \mathcal{L}(V, \mathbb{R}^k)$, and further they are isomorphisms, that is they are invertible, and so $\varphi_\beta \in \text{GL}(V, \mathbb{R}^k)$.

Proof: \ Linearity was shown in the previous theorem. To see that $\varphi_\beta$ is an isomorphism, note first that $\varphi_\beta$ takes bases to bases: if $\beta = (v_1, \ldots, v_k)$ is a basis for $V$, then $\varphi_\beta(v_i) = e_i$, since $v_i = 0v_1 + \cdots + 1v_i + \cdots + 0v_k$. Thus, it takes $\beta$ to the standard basis $\rho_k = (e_1, \ldots, e_k)$ for $\mathbb{R}^k$. Consequently, it is surjective, because if $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$, then

$$x = x_1e_1 + \cdots + x_ke_k = x_1\varphi_\beta(v_1) + \cdots + x_k\varphi_\beta(v_k) = \varphi_\beta(x_1v_1 + \cdots + x_kv_k)$$

If we let $v = x_1v_1 + \cdots + x_kv_k$, then we see that $v \in V$ satisfies $\varphi_\beta(v) = x$. But $\varphi_\beta$ is also injective: if $v \in \ker \varphi_\beta$, then $v = a_1v_1 + \cdots + a_kv_k$, so that

$$0 = \varphi_\beta(v) = \varphi_\beta(a_1v_1 + \cdots + a_kv_k) = a_1\varphi_\beta(v_1) + \cdots + a_k\varphi_\beta(v_k) = a_1e_1 + \cdots + a_ke_k$$

By the linear independence of the $e_i$ we must have $a_1 = \cdots = a_k = 0$, and so $v = 0$. Thus, $\varphi_\beta$ is also injective. \hfill \blacksquare

Theorem 0.23 \ Let $V$ and $W$ be finite-dimensional subspaces of $\mathbb{R}^n$ having ordered bases $\beta = (v_1, \ldots, v_k)$ and $\gamma = (w_1, \ldots, w_\ell)$, respectively, and let $T \in \mathcal{L}(V, W)$. Then for all $v \in V$ we have

$$[T(v)]_\gamma = [T]_\beta^\gamma[v]_\beta \quad (0.11)$$

In other words, if $D = [T]_\beta^\gamma$ is the matrix representation of $T$ in $\beta$ and $\gamma$ coordinates, with $T_D \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^\ell)$ the corresponding matrix multiplication map, if $A = [T]_\rho_\ell^\mu$ is the matrix representation of $T$ in standard coordinates, with corresponding $T_A \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^\ell)$, and if $\varphi_\beta \in \text{GL}(V, \mathbb{R}^k)$ and $\varphi_\gamma \in \text{GL}(W, \mathbb{R}^\ell)$ are the respective coordinate maps, with matrix representations $B^{-1}$ and $C^{-1}$, respectively, where $B = [v_1 \cdots v_k]$ and $C = [w_1 \cdots w_\ell]$, then

$$\varphi_\gamma \circ T = T_D \circ \varphi_\beta \quad \text{or, in terms of matrices,} \quad C^{-1}A = DB^{-1} \quad (0.12)$$

and the following diagrams commute:

\[
\begin{array}{ccc}
V & \xrightarrow{T} & W \\
\downarrow{\varphi_\beta} & & \downarrow{\varphi_\gamma} \\
\mathbb{R}^k & \xrightarrow{T_D} & \mathbb{R}^\ell
\end{array}
\quad \text{or, in terms of matrices,} \quad \begin{array}{ccc}
V & \xrightarrow{A} & W \\
\downarrow{B^{-1}} & & \downarrow{C^{-1}} \\
\mathbb{R}^k & \xrightarrow{D} & \mathbb{R}^\ell
\end{array}
\]

Proof: If $\beta = (v_1, \ldots, v_k)$ is an ordered basis for $V$ and $\gamma = (w_1, \ldots, w_\ell)$ is an ordered basis for $W$, then let

$$[T]_\beta = \begin{bmatrix} [T(v_1)]_\gamma & \cdots & [T(v_k)]_\gamma \end{bmatrix} = \begin{bmatrix} (a_{11}) & \cdots & (a_{1k}) \\
\vdots & \ddots & \vdots \\
(a_{\ell1}) & \cdots & (a_{\ellk}) \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\
\vdots & \ddots & \vdots \\
a_{\ell1} & \cdots & a_{\ellk} \end{bmatrix}$$

Now, for all $u \in V$ there are unique $b_1, \ldots, b_k \in \mathbb{R}$ such that $u = b_1v_1 + \cdots + b_kv_k$. Therefore,

$$T(u) = T(b_1v_1 + \cdots + b_kv_k) = b_1T(v_1) + \cdots + b_kT(v_k)$$
Indeed, the invertible matrix 

\[ P \]

so that by the linearity of \( \varphi \),

\[
[T(u)]_\gamma = \phi_\gamma(T(u)) = b_1 \phi_\gamma(T(v_1)) + \cdots + b_k \phi_\gamma(T(v_k))
\]

\[
= b_1 \phi_\gamma(T(v_1)) + \cdots + b_k \phi_\gamma(T(v_k))
\]

\[
= b_1[T(v_1)]_\gamma + \cdots + b_k[T(v_k)]_\gamma
\]

\[
= \left[ \begin{array}{c}
[T(v_1)]_\gamma \\
\vdots \\
[T(v_k)]_\gamma
\end{array} \right]
\]

\[
= \left[ \begin{array}{c}
b_1 \\
b_2 \\
\vdots \\
b_k
\end{array} \right]
\]

This shows that \( \varphi_\gamma \circ T = T_D \circ \varphi_\beta \), since \( [T(u)]_\gamma = (\varphi_\gamma \circ T)(u) \) and \( [T(u)]_\beta = (T_D \circ \varphi_\beta)(u) \). Finally, since \( \varphi_\beta(x) = B^{-1}x \) and \( \varphi_\gamma(y) = C^{-1}y \), we have the equivalent statement \( C^{-1}A = DB^{-1} \). ■

**Remark 0.24** Two square matrices \( A, B \in \mathbb{R}^{n \times n} \) are said to be similar, denoted \( A \sim B \), if there exists an invertible matrix \( P \in \text{GL}(n, \mathbb{R}) \) such that

\[ A = PBP^{-1} \]

Similarity is an equivalence relation (it is reflexive, symmetric and transitive):

1. (reflexivity) \( A \sim A \) because we can take \( P = I_n \), so that \( A = I_n A I_n^{-1} \)
2. (symmetry) If \( A \sim B \), then there is an invertible \( P \) such that \( A = PBP^{-1} \). But then left-multiplying by \( P^{-1} \) and right-multiplying \( P \) gives \( P^{-1}AP = B \), so since \( P^{-1} \) is invertible, we have \( B \sim A \).
3. (transitivity) If \( A \sim B \) and \( B \sim C \), then there are invertible matrices \( P \) and \( Q \) such that \( A = PBP^{-1} \) and \( B = QCQ^{-1} \). Plugging the second into the first gives

\[
A = PBP^{-1} = P(QCQ^{-1})P^{-1} = (PQ)C(Q^{-1}P^{-1}) = (PQ)C(PQ)^{-1}
\]

so since \( PQ \) is invertible, with inverse \( Q^{-1}P^{-1} \), \( A \sim C \).

In the previous theorem, if we take \( W = V \) and \( \gamma = \beta \), we’ll have \( B = C \), so that

\[
B^{-1}A = DB^{-1}, \quad \text{or} \quad A = BDB^{-1}
\]

which shows that \( A \sim D \), i.e.

\[
[T]_\rho \sim [T]_\beta
\]

Since similarity is an equivalence relation, if \( \beta \) and \( \gamma \) are any two bases for \( V \), then

\[
[T]_\beta \sim [T]_\rho \quad \text{and} \quad [T]_\rho \sim [T]_\gamma \implies [T]_\beta \sim [T]_\gamma
\]

This demonstrates that if we represent a linear transformation \( T \in \mathcal{L}(V, V) \) with respect to two different bases, then the corresponding matrices are similar

\[
[T]_\beta \sim [T]_\gamma
\]

Indeed, the invertible matrix \( P \) is \( B^{-1}C \), because \( [T]_\beta = B^{-1}[T]_\rho B \) and \( [T]_\rho = C[T]_\gamma C^{-1} \), so that

\[
[T]_\beta = B^{-1}[T]_\rho B = B^{-1}C[T]_\gamma C^{-1}B = (B^{-1}C)[T]_\gamma (B^{-1}C)^{-1}
\]

We will show below that the converse is also true: if two matrices are similar, then they represent the same linear transformation, possibly with respect to different bases! ■
Theorem 0.25  If \( V,W,Z \) are finite-dimensional subspaces of \( \mathbb{R}^n \), \( \mathbb{R}^m \) and \( \mathbb{R}^p \), respectively, with ordered bases \( \alpha, \beta \) and \( \gamma \), respectively, and if \( T \in \mathcal{L}(V,W) \) and \( U \in \mathcal{L}(W,Z) \), then

\[
[U \circ T]_\beta = [T]_\beta^\alpha [U]_\gamma^\beta
\]  

(0.13)

Proof: By the previous two theorems we have

\[
[U \circ T]_\alpha = [U]_\beta^\alpha \cdot [T]_\gamma^\beta = [U(T(v))]_\gamma = [U(T(v))]_\gamma
\]

which completes the proof.

\[\square\]

Corollary 0.26  If \( V \) is an \( k \)-dimensional subspace of \( \mathbb{R}^n \) with an ordered basis \( \beta \) and if \( I \in \mathcal{L}(V,V) \) is the identity operator, then \( [I]_\beta = I_k \in \mathbb{R}^{k \times k} \).

Proof: Given any \( T \in \mathcal{L}(V) \), we have \( T = I \circ T \), so that \( [T]_\beta = [I \circ T]_\beta = [I]_\beta [T]_\beta \), and similarly \( [T]_\beta = [T]_\beta [I]_\beta \). Hence, taking \( T = I \) we will have \( [I]_\beta^2 = [I]_\beta \). Note also that \( [I]_\beta^{-1} = [I]_\beta \), because \( [I]_\beta^{-1} [I]_\beta = [I]_\beta [I]_\beta^{-1} = [I \circ I^{-1}]_\beta = [I]_\beta \). But then \( [I]_\beta = I_n \), because any \( A \in \mathbb{R}^{n \times n} \) that is invertible and satisfies \( A^2 = A \) will satisfy \( A = I_n \).

\[\begin{align*}
A = AI_n &= (AA^{-1}) = (AA)A^{-1} = A^2A^{-1} = AA^{-1} = I_k \\
\text{An alternative proof follows directly from the definition, since} & \quad I(v_i) = v_i, \text{ so that} \quad [I(v_i)]_\beta = e_i, \text{ whence} \quad [I]_\beta = [I(v_1)]_\beta \cdots [I(v_k)]_\beta = [e_1 \cdots e_k] = I_k.
\end{align*}\]

\[\square\]

Theorem 0.27  If \( V \) and \( W \) are subspaces of \( \mathbb{R}^n \) with ordered bases \( \beta = (b_1, \ldots, b_k) \) and \( \gamma = (c_1, \ldots, c_\ell) \), respectively, then the function

\[
\Phi : \mathcal{L}(V,W) \to \mathbb{R}^{\ell \times k}
\]

\[
\Phi(T) = [T]_\beta^\gamma
\]

(0.14) (0.15)

is an isomorphism, that is \( \Phi \in \text{GL}(\mathcal{L}(V,W), \mathbb{R}^{\ell \times k}) \), and consequently the space of linear transformations is isomorphic to the space of all \( \ell \times k \) matrices:

\[\mathcal{L}(V,W) \cong \mathbb{R}^{\ell \times k}\]

(0.16)

Proof: First, \( \Phi \) is linear: there exist scalars \( r_{ij}, s_{ij} \in \mathbb{R} \), for \( i = 1, \ldots, \ell \) and \( j = 1, \ldots, k \), such that for any \( T \in \mathcal{L}(V,W) \) we have

\[
T(b_1) = r_{11} c_1 + \cdots + r_{1\ell} c_\ell \\
\vdots \\
T(b_k) = r_{k1} c_1 + \cdots + r_{k\ell} c_\ell
\]

\[
U(b_1) = s_{11} c_1 + \cdots + s_{1\ell} c_\ell \\
\vdots \\
U(b_k) = s_{k1} c_1 + \cdots + s_{k\ell} c_\ell
\]

Hence, for all \( s,t \in \mathbb{R} \) and \( j = 1, \ldots, n \) we have

\[
(sT + tU)(b_j) = sT(b_j) + tU(b_j) = s \sum_{i=1}^{\ell} r_{ij} c_i + t \sum_{i=1}^{\ell} s_{ij} c_i = \sum_{i=1}^{\ell} (sr_{ij} + ts_{ij}) c_i
\]
As a consequence of this and the rules of matrix addition and scalar multiplication we have

\[
\Phi(sT + tU) = [sT + tU]_\beta
\]

\[
= \begin{pmatrix}
   sr_{11} + ts_{11} & \cdots & sr_{1k} + ts_{1k} \\
   \vdots & \ddots & \vdots \\
   sr_{\ell1} + ts_{\ell1} & \cdots & sr_{\ell k} + ts_{\ell k}
\end{pmatrix}
\]

\[
= s \begin{pmatrix}
   r_{11} & \cdots & r_{1k} \\
   \vdots & \ddots & \vdots \\
   r_{\ell1} & \cdots & r_{\ell k}
\end{pmatrix} + t \begin{pmatrix}
   s_{11} & \cdots & s_{1k} \\
   \vdots & \ddots & \vdots \\
   s_{\ell1} & \cdots & s_{\ell k}
\end{pmatrix}
\]

\[
= s[T]_\beta^{\gamma} + t[U]_\beta^{\gamma}
\]

Moreover, \( \Phi \) is bijective, since for all \( T \in \mathcal{L}(V,W) \) there is only

\[
\{ \Phi(T) \}
\]

This makes \( \Phi \) onto, and also 1-1 because \( \ker(\Phi) = \{0\} \), the zero operator, because for \( O \in \mathcal{R}^{\ell \times k} \) there is only \( T \in \mathcal{L}(V,W) \) satisfying \( \Phi(T) = O \), because

\[
T(b_j) = 0c_1 + \cdots + 0c_\ell = 0
\]

defines a unique transformation, \( T_0 \).

**Theorem 0.28** Let \( V \) and \( W \) be subspaces of \( \mathbb{R}^n \) of the same dimension \( k \), with ordered bases \( \beta = (b_1, \ldots, b_k) \) and \( \gamma = (c_1, \ldots, c_k) \), respectively, and let \( T \in \mathcal{L}(V,W) \). Then \( T \) is an isomorphism iff \( [T]_\beta^{\gamma} \) is invertible, that is \( T \in \text{GL}(V,W) \), iff \( [T]_\gamma^{\beta} \in \text{GL}(k,\mathbb{R}) \). In this case

\[
[T^{-1}]_\gamma^{\beta} = ([T]_\beta^{\gamma})^{-1}
\]

**Proof:** If \( T \in \text{GL}(V,W) \) and \( \dim(V) = \dim(W) = k \), then \( T^{-1} \in \mathcal{L}(W,V) \), and \( T \circ T^{-1} = I_W \) and \( T^{-1} \circ T = I_V \), so that by Theorem 0.25 and Corollary 0.26 we have

\[
[T]_\beta^{[T^{-1}]_\gamma^{\beta}[T]_\beta^{\gamma} = [T \circ T^{-1}]_\gamma = [I_W]_\gamma = [I_V]_\beta = [T^{-1} \circ T]_\beta = [T^{-1}]_\gamma^{\beta}[T]_\beta^{\gamma}
\]

so that \( [T]_\beta^{\gamma} \) is invertible with inverse \( [T^{-1}]_\gamma^{\beta} \), and by the uniqueness of the multiplicative inverse in \( \mathbb{R}^{k \times k} \), which follows from the uniqueness of the inverse of \( T^{-1} \in \text{GL}(\mathbb{R}^k, \mathbb{R}^k) \), we have

\[
[T^{-1}]_\gamma^{\beta} = ([T]_\beta^{\gamma})^{-1}
\]

Conversely, if \( A = [T]_\beta^{\gamma} \) is invertible, there is a \( n \times n \) matrix \( B \) such that \( AB = BA = I_n \). Define \( U \in \mathcal{L}(W,V) \) on the basis elements as follows, \( U(c_j) = v_j = \sum_{i=1}^n b_{ij}b_{ij} \), and extend \( U \) by linearity. Then \( B = [U]_\gamma^{\beta} \). To show that \( U = T^{-1} \), note that

\[
[U \circ T]_\beta^{\gamma} = [U]_\gamma^{\beta}[T]_\beta^{\gamma} = BA = I_n = [I_V]_\beta \quad \text{and} \quad [T \circ U]_\gamma = [T]_\gamma^{\beta}[U]_\gamma^{\beta} = AB = I_n = [I_W]_\gamma
\]

But since \( \Phi \in \text{GL}(\mathcal{L}(V,W), \mathcal{R}^{k \times k}) \) is an isomorphism, and therefore 1-1, we must have that

\[
U \circ T = I_V \quad \text{and} \quad T \circ U = I_W
\]

By the uniqueness of the inverse, however, \( U = T^{-1} \), and \( T \) is an isomorphism.
0.3 Change of Coordinates

0.3.1 Definitions

We now define the change of coordinates, or change of basis, operator. If $V$ is a $k$-dimensional subspace of $\mathbb{R}^n$ and $\beta = (b_1, \ldots, b_k)$ and $\gamma = (c_1, \ldots, c_k)$ are two ordered bases for $V$, then the coordinate maps $\phi_\beta, \phi_\gamma \in \text{GL}(V, \mathbb{R}^k)$, which are isomorphisms by Corollary 0.22 above, may be used to define a change of coordinates operator $\phi_{\beta,\gamma} \in \text{GL}(\mathbb{R}^k, \mathbb{R}^k)$ changing $\beta$ coordinates into $\gamma$ coordinates, that is having the property

$$\phi_{\beta,\gamma}(\mathbf{v}_\beta) = \mathbf{v}_\gamma \quad (0.18)$$

We define the operator as follows:

$$\phi_{\beta,\gamma} := \phi_\gamma \circ \phi_\beta^{-1} \quad (0.19)$$

The relationship between these three functions is illustrated in the following commutative diagram:

$$\begin{array}{ccc}
\mathbb{R}^k & \xrightarrow{\phi_\gamma \circ \phi_\beta^{-1}} & \mathbb{R}^k \\
\phi_\beta \downarrow & & \downarrow \phi_\gamma \\
V & & V
\end{array}$$

As we will see below, the change of coordinates operator has a matrix representation

$$M_{\beta,\gamma} = [\phi_{\beta,\gamma}]_\rho = [\phi_\gamma]_\beta^\rho = \left[ \begin{array}{cccc} [b_1]_\gamma & [b_2]_\gamma & \cdots & [b_k]_\gamma \end{array} \right] \quad (0.20)$$

0.3.2 Properties of Change-of-Coordinate Maps and Matrices

**Theorem 0.29 (Change of Coordinate Matrix)** Let $V$ be a $k$-dimensional subspace of $\mathbb{R}^n$ and let $\beta = (b_1, \ldots, b_k)$ and $\gamma = (c_1, \ldots, c_k)$ be two ordered bases for $V$. Since $\phi_\beta$ and $\phi_\gamma$ are isomorphisms, the following diagram commutes,

$$\begin{array}{ccc}
\mathbb{R}^k & \xrightarrow{\phi_\gamma \circ \phi_\beta^{-1}} & \mathbb{R}^k \\
\phi_\beta \downarrow & & \downarrow \phi_\gamma \\
V & & V
\end{array}$$

and the change of basis operator $\phi_{\beta,\gamma} := \phi_\gamma \circ \phi_\beta^{-1} \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^k)$, changing $\beta$ coordinates into $\gamma$ coordinates, is an isomorphism. It’s matrix representation,

$$M_{\beta,\gamma} = [\phi_{\beta,\gamma}]_\rho \in \mathbb{R}^{k \times k} \quad (0.21)$$

where $\rho = (e_1, \ldots, e_n)$ is the standard ordered basis for $\mathbb{R}^k$, is called the change of coordinate matrix, and it satisfies the following conditions:

1. $M_{\beta,\gamma} = \left[ \phi_{\beta,\gamma} \right]_\rho = \left[ \phi_\gamma \right]_\beta^\rho = \left[ \begin{array}{cccc} [b_1]_\gamma & [b_2]_\gamma & \cdots & [b_k]_\gamma \end{array} \right]$

2. $[\mathbf{v}]_\gamma = M_{\beta,\gamma} [\mathbf{v}]_\beta$, $\forall \mathbf{v} \in V$

3. $M_{\beta,\gamma}$ is invertible and $M_{\beta,\gamma}^{-1} = M_{\gamma,\beta} = \left[ \phi_\beta \right]_\gamma^\rho$
Proof: The first point is shown as follows:

\[ M_{\beta,\gamma} = [\phi_{\beta,\gamma}]_\rho = \begin{bmatrix} \phi_{\beta,\gamma}(e_1) & \phi_{\beta,\gamma}(e_2) & \cdots & \phi_{\beta,\gamma}(e_k) \end{bmatrix} \]

\[ = \begin{bmatrix} (\phi_\gamma \circ \phi_\beta^{-1})(e_1) & (\phi_\gamma \circ \phi_\beta^{-1})(e_2) & \cdots & (\phi_\gamma \circ \phi_\beta^{-1})(e_k) \end{bmatrix} \]

\[ = \begin{bmatrix} (\phi_\gamma \circ \phi_\beta^{-1})(b_1|\beta) & (\phi_\gamma \circ \phi_\beta^{-1})(b_2|\beta) & \cdots & (\phi_\gamma \circ \phi_\beta^{-1})(b_k|\beta) \end{bmatrix} \]

\[ = \begin{bmatrix} \phi_{\gamma}(b_1) & \phi_{\gamma}(b_2) & \cdots & \phi_{\gamma}(b_k) \end{bmatrix} \]

\[ = \begin{bmatrix} b_1|\gamma & b_2|\gamma & \cdots & b_n|\gamma \end{bmatrix} \]

or, alternatively, by Theorem 0.25 and Theorem 0.28 we have that

\[ M_{\beta,\gamma} = [\phi_{\beta,\gamma}]_\rho = [\phi_\gamma \circ \phi_\beta^{-1}]_\rho = [\phi_\gamma]_\rho[\phi_\beta^{-1}]_\rho = [\phi_\gamma]_\rho[I_n] = [\phi_\gamma]_\rho = M_{\beta,\gamma} \]

The second point follows from Theorem 0.23, since

\[ \phi_\gamma(v) = (\phi_\gamma \circ I)(v) = (\phi_\gamma \circ (\phi_\beta^{-1} \circ \phi_\beta))(v) = ((\phi_\gamma \circ \phi_\beta^{-1}) \circ \phi_\beta)(v) \]

implies that

\[ [v]_\gamma = [\phi_\gamma(v)]_\rho = [((\phi_\gamma \circ \phi_\beta^{-1}) \circ \phi_\beta)(v)]_\rho = [\phi_\gamma \circ \phi_\beta^{-1}]_\rho[\phi_\beta(v)]_\rho = [\phi_{\beta,\gamma}]_\rho[v]_\beta = M_{\beta,\gamma}[v]_\beta \]

And the last point follows from the fact that \( \phi_\beta \) and \( \phi_\gamma \) are isomorphism, so that \( \phi_{\beta,\gamma} \) is an isomorphism, and hence \( \phi_{\beta,\gamma}^{-1} \in \mathcal{L}(\mathbb{R}^k) \) is an isomorphism, and because the diagram above commutes we must have

\[ \phi_{\beta,\gamma}^{-1} = (\phi_\gamma \circ \phi_\beta^{-1})^{-1} = \phi_\beta \circ \phi_\gamma^{-1} = \phi_{\beta,\gamma} \]

so that by (1)

\[ M_{\beta,\gamma}^{-1} = [\phi_{\beta,\gamma}]_\rho = [\phi_{\gamma,\beta}]_\rho = M_{\gamma,\beta} \]

or alternatively by Theorem 0.28

\[ M_{\beta,\gamma}^{-1} = ([\phi_{\beta,\gamma}]_\rho)^{-1} = [\phi_{\gamma,\beta}]_\rho = [\phi_\beta]_\rho = M_{\gamma,\beta} \]

\[ \square \]

Corollary 0.30 (Change of Basis) Let \( V \) and \( W \) be subspaces of \( \mathbb{R}^n \) and let \( (\beta, \gamma) \) and \( (\beta', \gamma') \) be pairs of ordered bases for \( V \) and \( W \), respectively. If \( T \in \mathcal{L}(V, W) \), then

\[
[T]_{\beta'}^{\gamma'} = M_{\gamma',\gamma}[T]_{\beta}^{\gamma} M_{\beta',\beta}
\]

(0.22)

\[
M_{\gamma',\gamma}[T]_{\beta}^{\gamma} M_{\beta',\beta}^{-1}
\]

(0.23)

where \( M_{\gamma',\gamma} \) and \( M_{\beta',\beta} \) are change of coordinate matrices. That is, the following diagram commutes.
Proof: This follows from the fact that if $\beta = (b_1, \ldots, b_k)$, $\beta' = (b'_1, \ldots, b'_k)$, $\gamma = (c_1, \ldots, c_\ell)$, $\gamma' = (c'_1, \ldots, c'_\ell)$, then for each $i = 1, \ldots, k$ we have

$$\left[T(b'_i)\right]_{\gamma'} = \left[(\phi_{\gamma,\gamma'} \circ T \circ \phi_{\beta,\beta'}^{-1})(b'_i)\right]_{\gamma'}$$

so

$$\left[T\right]_{\beta'} = \left[\phi_{\gamma,\gamma'} \circ T \circ \phi_{\beta,\beta'}^{-1}\right]_{\beta'}$$

$$= \left[\phi_{\gamma,\gamma'}\right]_{\rho_n} \left[T\right]_{\beta} \left[\phi_{\beta,\beta'}^{-1}\right]_{\rho_n}$$

$$= \left[\phi_{\gamma,\gamma'}\right]_{\rho_n} \left[T\right]_{\beta} \left(\left[\phi_{\beta,\beta'}\right]_{\rho_n}\right)^{-1}$$

$$= M_{\gamma,\gamma'} \left[T\right]_{\beta} M_{\beta,\beta'}^{-1}$$

which completes the proof. \[\blacksquare\]

Corollary 0.31 (Change of Basis for a Linear Operator) If $V$ is a subspace of $\mathbb{R}^n$ with ordered bases $\beta$ and $\gamma$, and $T \in L(V)$, then

$$\left[T\right]_{\gamma} = M_{\beta,\gamma} \left[T\right]_{\beta} M_{\beta,\gamma}^{-1} \tag{0.24}$$

where $M_{\beta,\gamma}$ is the change of coordinates matrix. \[\blacksquare\]

Corollary 0.32 If we are given any two of the following:

1. $A \in \mathbb{R}^n$ invertible
2. an ordered basis $\beta$ for $\mathbb{R}^n$
3. an ordered basis $\gamma$ for $\mathbb{R}^n$

The third is uniquely determined by the equation $A = M_{\beta,\gamma}$, where $M_{\beta,\gamma}$ is the change of coordinates matrix of the previous theorem.

Proof: If we have $A = M_{\beta,\gamma} = \left[\phi_\beta\right]_{\gamma} = \left[b_1\right]_{\gamma} \left[b_2\right]_{\gamma} \cdots \left[b_n\right]_{\gamma}$, suppose we know $A$ and $\gamma$. Then, $\left[b_i\right]_{\gamma}$ is given by $A$, so $b_i = A_{i1}c_1 + \cdots + A_{in}c_n$, so $\beta$ is uniquely determined. If $\beta$ and $\gamma$ are given, then by the previous theorem $M_{\beta,\gamma}$ is given by $M_{\beta,\gamma} = \left[b_1\right]_{\gamma} \left[b_2\right]_{\gamma} \cdots \left[b_n\right]_{\gamma}$. Lastly, if $A$ and $\beta$ are given, then $\gamma$ is given by the first case applied to $A^{-1} = M_{\gamma,\beta}$. \[\blacksquare\]