Matrix Representation of Linear Maps

Math 422

For simplicity we’ll work in the plane \( \mathbb{R}^2 \), although much of this material extends directly to general vector spaces. Recall that a basis for \( \mathbb{R}^2 \) consists of two non-zero non-parallel vectors. Choose an ordered basis \( \mathcal{B} = \{ \mathbf{u}, \mathbf{v} \} \) for \( \mathbb{R}^2 \) and recall that every vector \( \mathbf{w} \in \mathbb{R}^2 \) can be written uniquely as a linear combination of elements in \( \mathcal{B} \), i.e.,

\[
\mathbf{w} = s \mathbf{u} + t \mathbf{v}
\]

for some \( s, t \in \mathbb{R} \). These unique coefficients \( s \) and \( t \) are referred to as the \( \mathcal{B} \)-coordinates of \( \mathbf{w} \) and we write

\[
[\mathbf{w}]_\mathcal{B} = \begin{bmatrix} s \\ t \end{bmatrix}.
\]

[\mathbf{w}]_\mathcal{B} is called the \( \mathcal{B} \)-coordinate matrix for \( \mathbf{w} \).

**Definition 1** A map \( T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \) is linear iff

\[
T(ax + by) = aT(x) + bT(y),
\]

for all \( x, y \in \mathbb{R}^2 \) and all \( a, b \in \mathbb{R} \).

Let \( T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \) be a linear map. Apply \( T \) to each basis vector and write \( T(\mathbf{u}) \) and \( T(\mathbf{v}) \) in the basis \( \mathcal{B} \) as

\[
T(\mathbf{u}) = a\mathbf{u} + c\mathbf{v} \quad \text{and} \quad T(\mathbf{v}) = b\mathbf{u} + d\mathbf{v};
\]

then

\[
[T(\mathbf{u})]_\mathcal{B} = \begin{bmatrix} a \\ c \end{bmatrix} \quad \text{and} \quad [T(\mathbf{v})]_\mathcal{B} = \begin{bmatrix} b \\ d \end{bmatrix}.
\]

For a general vector \( \mathbf{x} = x_1 \mathbf{u} + x_2 \mathbf{v} \) we have

\[
T(\mathbf{x}) = T(x_1 \mathbf{u} + x_2 \mathbf{v}) = x_1 T(\mathbf{u}) + x_2 T(\mathbf{v}) = x_1 (a\mathbf{u} + c\mathbf{v}) + x_2 (b\mathbf{u} + d\mathbf{v})
\]

\[
= x_1 a\mathbf{u} + x_1 c\mathbf{v} + x_2 b\mathbf{u} + x_2 d\mathbf{v} = (x_1 a + x_2 b)\mathbf{u} + (x_1 c + x_2 d)\mathbf{v}.
\]

Thus

\[
[T(\mathbf{x})]_\mathcal{B} = [(x_1 a + x_2 b)\mathbf{u} + (x_1 c + x_2 d)\mathbf{v}]_\mathcal{B} = \begin{bmatrix} x_1 a + x_2 b \\ x_1 c + x_2 d \end{bmatrix}
\]

\[
= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [T(\mathbf{u})]_\mathcal{B} \cdot [T(\mathbf{v})]_\mathcal{B} \cdot [\mathbf{x}]_\mathcal{B}
\]

The matrix for \( T \) in the basis \( \mathcal{B} \) is:

\[
[T]_\mathcal{B} = [[T(\mathbf{u})]_\mathcal{B} \mid [T(\mathbf{v})]_\mathcal{B}] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}
\]

and in this notation we have

\[
[T]_\mathcal{B} \cdot [\mathbf{x}]_\mathcal{B} = [T(\mathbf{x})]_\mathcal{B}. \tag{1}
\]

Let \( [\mathbf{x}] \) denote the coordinate matrix for \( \mathbf{x} \) in the standard basis \( \{ \mathbf{e}_1, \mathbf{e}_2 \} \) and write

\[
[\mathbf{u}] = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{and} \quad [\mathbf{v}] = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.
\]

Consider the matrix

\[
P = [[\mathbf{u}] \mid [\mathbf{v}]] = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}.
\]

Multiplication by \( P \) changes the coordinate matrix for a vector \( \mathbf{x} \) from one in \( \mathcal{B} \)-coordinates to one in standard coordinates:

\[
P \cdot [\mathbf{x}]_\mathcal{B} = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 u_1 + x_2 v_1 \\ x_1 u_2 + x_2 v_2 \end{bmatrix} = [\mathbf{x}].
\]
$P$ is called the \textit{transition matrix from $B$ to standard coordinates}. Since $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, \(\det(P) \neq 0\) and $P$ is invertible. Multiplying both sides on the left by $P^{-1}$ gives

\[
[x]_B = P^{-1} [x].
\] (2)

Thus, multiplication by $P^{-1}$ changes the coordinate matrix for $x$ from one in standard coordinates to one in $B$-coordinates.

Combining (1) and (2) we have

\[
[T]_B P^{-1} [x] = [T]_B [x]_B = [T(x)]_B = P^{-1} [T(x)].
\]

or equivalently

\[
P[T]_B P^{-1} [x] = [T] [x].
\]

This says that the matrix for $T$ in the standard basis is

\[
[T] = P[T]_B P^{-1}.
\]

Since $\det(P^{-1}) \det(P) = \det(P^{-1}P) = \det(I) = 1$ we have

\[
\det [T] = \det(P^{-1} [T]_B P) = \det(P^{-1}) \det [T]_B \det(P) = \det [T]_B.
\]

We have proved

\textbf{Theorem 2} Let $B$ be a basis for $\mathbb{R}^2$ and let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear map. Then

\[
\det [T]_B = \det [T],
\]

\textit{i.e., the determinant of the matrix for $T$ is independent of the choice of basis.}

It makes sense, therefore, to talk about the “determinant” of a linear map.

\textbf{Definition 3} Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear map. Then the \textit{determinant of $T$} is defined by

\[
\det(T) = \det [T].
\]

The map $T$ is said to be \textit{non-singular} whenever $\det(T) \neq 0$.

\textbf{Example 4} Let $B = \{\mathbf{u}, \mathbf{v}\}$ be the ordered basis for $\mathbb{R}^2$ with $[\mathbf{u}] = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $[\mathbf{v}] = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Define a linear map $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by

\[
T(\mathbf{u}) = 3\mathbf{u} - 4\mathbf{v} \text{ and } T(\mathbf{v}) = -\mathbf{u} + 3\mathbf{v}.
\]

Then

\[
[T]_B = \begin{bmatrix} 3 & -1 \\ -4 & 3 \end{bmatrix}.
\]

The transition matrix $P$ from $B$ to standard coordinates is

\[
P = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}
\]

and consequently

\[
P^{-1} = \frac{1}{5} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}.
\]

Therefore the matrix for $T$ in the standard basis is

\[
[T] = P[T]_B P^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -4 & 3 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 5 \end{bmatrix}.
\]

Observe that $\det [T] = \det [T]_B = 5$. 

2
Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a non-singular linear map and let $b \in \mathbb{R}^2$. If $x$ is unknown, then $[T]$ is coefficient matrix for a linear system

$$[T]x = b.$$  \hfill (3)

Since $[T]$ is invertible, equation (3) has the (unique) solution

$$x = [T]^{-1}b,$$

which is to say that the map $T$ is surjective. Furthermore, if $[T]x = [T]y$, then multiplying on the left by $[T]^{-1}$ gives $x = y$ and the map $T$ is also injective. We have proved

**Theorem 5** Every non-singular linear map is bijective.

**Definition 6** A bijective linear map $T : V \to W$ is called an isomorphism of vector spaces. When $T : V \to W$ is an isomorphism, the spaces $V$ and $W$ are said to be isomorphic.