Change of Basis of a Linear Transformation

For some transformations, it may be easier to work with a coordinate system other than the standard \( x_1 - x_2 \) coordinate system. For example, in the case where the transformation is a projection onto a line in \( \mathbb{R}^2 \), it might be easier to work with a coordinate system where one of the axes is on the line and the other is perpendicular to the line.

This change of coordinate systems is equivalent to changing the basis from the standard basis to some other \( \beta \) basis. For these types of problems, we want to know how to find the matrix of the linear transformation with respect to the new \( \beta \) basis.

Example

Let \( L \) be the line in \( \mathbb{R}^2 \) spanned by the vector \(
\begin{bmatrix}
2
1
\end{bmatrix}
\). Let \( T \) be the linear transformation from \( \mathbb{R}^2 \) to \( \mathbb{R}^2 \) that projects any vector orthogonally onto \( L \), as shown below.
The standard matrix of this linear transformation is:

\[ A = \frac{1}{5} \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \]

In this case, it will help to introduce a coordinate system where one of the axes, call it \( c_1 \), is along the line \( L \) and the other axis, call it \( c_2 \), is perpendicular to \( L \).

If we use this new \( c_1 - c_2 \) coordinate system, then \( T \) transforms any vector

\[ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \]

into

\[ \begin{bmatrix} c_1 \\ 0 \end{bmatrix} \]

In \( c_1 - c_2 \) coordinates, \( T \) is given by the matrix \( B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \), since

\[ \begin{bmatrix} c_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \]

To be more precise, start by introducing a basis \( \beta = \{\vec{v}_1, \vec{v}_2\} \) of \( \mathbb{R}^2 \) with \( \vec{v}_1 \) on \( L \) and \( \vec{v}_2 \) perpendicular to \( L \). For example, \( \beta = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\} \).

If \( [\vec{x}]_\beta = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \), then \( [T(\vec{x})]_\beta = \begin{bmatrix} c_1 \\ 0 \end{bmatrix} \) (as illustrated on the next page).

Matrix \( B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \) transforms \( [\vec{x}]_\beta \) into \( [T(\vec{x})]_\beta \) and is called the \( \beta \)-matrix of \( T \):

\[ [T(\vec{x})]_\beta = B [\vec{x}]_\beta \]
In Summary: Consider a linear transformation $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ and a basis $\beta$ of of $\mathbb{R}^n$ consisting of the $n$ vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$, each $\vec{v}_i \in \mathbb{R}^n$. The $n \times n$ matrix $B$ that transforms $[\vec{x}]_\beta$ into $[T(\vec{x})]_\beta$ is called the $\beta$-matrix of $T$:

$$[T(\vec{x})]_\beta = B [\vec{x}]_\beta$$

How can we construct the $\beta$-matrix $B$ of a linear transformation $T$ column by column?

$$\vec{x} = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n,$$

then

$$T(\vec{x}) = c_1 T(\vec{v}_1) + \cdots + c_n T(\vec{v}_n),$$

by linearity of $T$, and

$$[T(\vec{x})]_\beta = c_1 [T(\vec{v}_1)]_\beta + \cdots + c_n [T(\vec{v}_n)]_\beta = \begin{bmatrix} [T(\vec{v}_1)]_\beta & \cdots & [T(\vec{v}_n)]_\beta \end{bmatrix} [\vec{x}]_\beta$$

In Summary: Consider a linear transformation $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ and a basis $\beta$ of of $\mathbb{R}^n$ consisting of the $n$ vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$, each $\vec{v}_i \in \mathbb{R}^n$. The $\beta$-matrix of $T$ is

$$B = \begin{bmatrix} [T(\vec{v}_1)]_\beta & \cdots & [T(\vec{v}_n)]_\beta \end{bmatrix};$$

that is, the columns of $B$ are the $\beta$-coordinate vectors of $T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_n)$.
If we are given the standard matrix $A$ of a linear transformation (matrix with respect to the standard basis) and a basis $\beta$, how can we find the $\beta$-matrix directly from $A$ and $\beta$?

Consider a linear transformation $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ and a basis $\beta$ of $\mathbb{R}^n$ consisting of the $n$ vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$, each $\vec{v}_i \in \mathbb{R}^n$.

Recall that the $\beta$-coordinate vector of any vector $\vec{x}$ can be found by:

$$\vec{x} = S [\vec{x}]_{\beta}, \quad \text{where} \quad S = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$$

Also recall that the output of a linear transformation is a vector. Therefore, just as we found the $\beta$-coordinate vector of any $\vec{x}$, we can also find the $\beta$-coordinate vector of any $T(\vec{x})$, or

$$T(\vec{x}) = S [T(\vec{x})]_{\beta}$$

(i.e., the $\beta$-coordinates of $T(\vec{x})$ are found by solving $T(\vec{x}) = S [T(\vec{x})]_{\beta}$)

Now consider the following diagram:

Given $A$ and $S$, how could we get directly from $[\vec{x}]_{\beta}$ to $[T(\vec{x})]_{\beta}$, or, how to find $B$?
If we are given a basis $\beta$ of $\mathbb{R}^n$ and a linear transformation $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, then we know the $S$ and $A$ matrices.

So the problem is, how can we find $B$ given $S$ and $A$?

Consider the two paths from $[\bar{x}]_\beta$ to $T(\bar{x})$:

1. Going from $[\bar{x}]_\beta$ to $T(\bar{x})$ through $\bar{x}$ (clockwise in diagram starting at lower left)

   $T(\bar{x}) = A\bar{x} = A(S[\bar{x}]_\beta) = AS[\bar{x}]_\beta \quad (1)$

2. Going from $[\bar{x}]_\beta$ to $T(\bar{x})$ through $[T(\bar{x})]_\beta$ (counterclockwise in diagram starting at lower left)

   $T(\bar{x}) = S[T(\bar{x})]_\beta = S(B[\bar{x}]_\beta) = SB[\bar{x}]_\beta \quad (2)$

Equating (1) and (2) gives

$T(\bar{x}) = AS[\bar{x}]_\beta = SB[\bar{x}]_\beta$

$AS = SB$

$S^{-1}AS = S^{-1}SB$

$S^{-1}AS = B$

$A$ is called the “standard matrix” of $T$ ($T$ using the standard basis)

$B$ is called the “$\beta$-matrix” of $T$ ($T$ using the $\beta$ basis)

$B = S^{-1}AS$, where $S = \begin{bmatrix} \bar{v}_1 & \cdots & \bar{v}_n \end{bmatrix}$

$(\beta = \{\bar{v}_1, \ldots, \bar{v}_n\}$ basis of $\mathbb{R}^n)$

$n \times n$
Conversely, we can also find an equation for $A$ given $B$:

\[ S^{-1}AS = B \]

\[ SS^{-1}AS = SB \]

\[ AS = SB \]

\[ ASS^{-1} = SBS^{-1} \]

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

To summarize:

If we are given the standard matrix, $A$, of a linear transformation $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, and a basis $\beta$ of $\mathbb{R}^n$, then we can find the $\beta$-matrix of $T$ by the equation

\[ B = S^{-1}AS \]

If we are given a $\beta$-matrix, $B$, of a linear transformation $T$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, where $\beta$ is a basis of $\mathbb{R}^n$, then we can find the standard matrix of $T$ by the equation

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

Corollary: If there is an invertible matrix $S$ (remember, any matrix formed by a set of basis vectors will be invertible) such that

\[ AS = SB, \quad \text{or} \quad B = S^{-1}AS, \]

then we say that $A$ is similar to $B$. 