Linear transformation and the change of basis

Abstract

This short note supplements the Linear algebra part of ZCA 110. In particular it discusses in understandable language (i) the idea of linear transformation involving different bases, as discussed in Chapter 12, Ayers, and (i) the idea of bases and coordinates, page 88-89, Ayers.

1 Going from one basis to another

Consider a generic $n-$vector, X, in $V_n(R)$. The vector X can be represented in different basis (as a simile: think of the appearance of an actor viewed through different coloured glasses by different autdience. Despite it is the same actor, to different audience the actor appears differently.) As an illustration, we will discuss how the vector be represented in two different basis. Let's agree to call these two generic basis the $W-$basis and the $Z-$basis.

The W-basis consists of a set of n $n-$vector, namely $\{W_1, W_2, \cdots, W_n\}$, where each of the W_i, $i = 1, 2, \cdots n$ is an n by 1 column vector, $W_i = (w_1, w_2, \cdots, w_n)^T$. Similarly, the $Z-$basis consists of the set $\{Z_1, Z_2, \cdots, Z_n\}$, $Z_i = (z_1, z_2, \cdots, z_n)^T$, $i = 1, 2, \cdots n$. The connection between the two bases can be worked out via the following consideration:

In the $W-$basis, the vector X presents itself as a linear combination in terms of W_i's, i.e.

$$X = a_1 W_1 + a_2 W_2 + \cdots a_n W_n. \tag{1}$$

a_i are scalars called the components of X in the $W-$basis. By definition, the components vector of X in the $W-$ basis is the column vector that contains all of the components, or coordinate, of vector X in the $W-$basis. It is denoted by $X_W = (a_1, a_2, \cdots, a_n)^T$. Let us arrange all of the basis vector W_i column-by-colum into the matrix $W = (W_1, W_2, \cdots, W_n)$, where W is a $n \times n$ matrix. Now, Eq. (1) can be compactly written in the form of

$$X = a_1 W_1 + a_2 W_2 + \cdots a_n W_n = (W_1, W_2, W_3, \cdots, W_n)(a_1, a_2, \cdots, a_n)^T = W \cdot X_W. \tag{2}$$

X_W is the coordinate vector of X relative to the W-basis.

Similarly, if the vector X were to be represented in the $Z-$basis,

$$X = b_1 Z_1 + b_2 Z_2 + \cdots b_n Z_n = (Z_1, Z_2, Z_3, \cdots, Z_n)(b_1, b_2, \cdots, b_n)^T = Z \cdot X_Z. \tag{3}$$

$X_Z = (b_1, b_2, \cdots, b_n)^T$ is the coordinate vector of X relative to the W-basis.

The vector X is the same vector irrespective of its basis representation, hence

$$X \quad = \quad W \cdot X_W = Z \cdot X_Z \tag{4}$$

Eq. (4) relates the coordinate vector of vector X represented in the Z-basis to that in the W-basis. The coodinate vector in one basis can be determined if the coordinate vector in the other is known, and vice versa.

For example, if we know X_W, we can determine X_Z by making use of Eq. (4): We form the matrix

$$P = Z^{-1} \cdot W, \tag{5}$$

operate it to X_W from the left to obtain

$$X_Z = P \cdot X_W = (Z^{-1} \cdot W) \cdot X_W. \tag{6}$$

Conversely, we can obtain X_W if X_Z is known via

$$X_W = P^{-1} X_Z. \tag{7}$$

2 Linear Transformation

In simple language, a transformation is an operation that operates onto a vector to make it into another vector. Normally the operation is realised via matrix multiplication. Say X is a vector to be transformed into anoter vector, call it Y. To implement the transformation, we will operate a matrix A onto X to make it into Y. Symbolically, $X \xrightarrow{A} Y$; operationally, $Y = AX$. The transformation matrix A contains the information (instruction) of how the vector is to be transformed (e.g. to rotate X about the origin by +90 degree, to reflect the vector X about the origin, etc.). Y, the resultant vector under the transformation, is called the image of X under transformation A.

A transformation can be carried out in any basis. Consider a vector X is tranformed into vector Y. Such a transformation can be represented in both the W-basis and the Z-basis. In each basis the transformation takes on different forms. Say A is the transformation matrix in the W-basis representation, whereas B is the correspoinding transformation in the Z-basis representation. The linear transformations in both bases are given by:

$$\begin{array}{ll}
W\text{-basis} & Z\text{-basis} \\
 & \\
W = (W_1, W_2, \cdots, W_n) & Z = (Z_1, Z_2, \cdots, Z_n) \\
X \xrightarrow{A} Y & X \xrightarrow{B} Y \\
 & \\
\text{In component form} & \text{In component form} \\
X_W \xrightarrow{A} Y_W & X_Z \xrightarrow{B} Y_Z \\
 & \\
\text{In matrix form} & \text{In matrix form} \\
Y_W = AX_W & Y_Z = BX_Z
\end{array}$$

Question: What is the definition of 'linear transformation'? (*Hint:* Refer to Ayers, page 94.)

Now, we shall prove that: If

$$\begin{aligned}
X_W &\xrightarrow{A} Y_W \text{ in the } W\text{-basis} \\
X_Z &\xrightarrow{B} Y_Z \text{ in the } Z\text{-basis,}
\end{aligned}$$

then the two transformation A, B are similar, i.e.,

$$B = Q^{-1}AQ,$$

where $Q = P^{-1} = (Z^{-1}W)^{-1} = W^{-1}Z$.

Question: What does it mean, mathematically, when it is said that matrix A and matrix B are 'similar'?

The proof is as followed: We begin with

$$Y_Z = BX_Z. \tag{8}$$

In Eq. (8), the LHS, i.e. Y_Z is related to Y_W via $Y_Z = PY_W$ [see Eq. (6)], whereas X_Z in the RHS is related to X_W via $X_Z = PX_W$ [see Eq. (7)]. Hence, Eq. (8) can be written as

$$\begin{aligned}
& PY_W = B(PX_W) \\
\Rightarrow \quad & Y_W = (P^{-1}BP)X_W. \qquad (9)
\end{aligned}$$

Eq. (9) is just the tranformation of X_W into Y_W by A (in the W-basis), i.e.

$$Y_W = (P^{-1}BP)X_W \equiv AX_W. \tag{10}$$

Hence, we can identify

$$A = P^{-1}BP,$$

or

$$B = PAP^{-1} = Q^{-1}AQ,$$

where P is given by Eq. (5).

Questions: (*i*) Attempt example 2, page 96, Ayers, yourself. (*ii*) Attempt Solved Problem 1, in the same page, yourself. Try to make yourself proud by solving these problems without reading the solutions.