Change of basis with commutative diagrams

July 20, 2021

1 Column vectors and vector spaces

We often identify elements of $\mathbb{R}^n$, and more generally elements of n-dimensional vector spaces, with $n \times 1$ matrices, the so-called column vectors. In doing so we make a, often implicit, choice of basis. This is usually harmless, but in this particular topic it will pay off to examine more closely this natural identification. Throughout the rest of this article we let V be a real finite-dimensional vector space; and we suppose $B = \{b_1, \dots, b_n\}$ is an ordered basis of V—here "ordered" just means that we care about the order in which we write the basis vectors.

We want a correspondence between V and $M_{n\times 1}(\mathbb{R})$, the latter being the set of $n \times 1$ real matrices. To achieve this we define a function $\tau_B \colon V \to M_{n\times 1}(\mathbb{R})$ in the following way. For each $v \in V$ there are unique scalars $c_1, \dots c_n \in \mathbb{R}$ such that $v = c_1 b_1 + \ldots + c_n b_n$. Then we let

$$\tau_B(v) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Clearly this representation is highly dependant on the choice of basis B we make. If V is $\mathbb{R}^n$ for example, and we choose the standard basis denoted by S then

$$\tau_S(c_1, \dots, c_n) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

but this is, in general, not the case for arbitrary bases.

It is worth pointing out that τ_B is a genuine bijection since it has an inverse $\tau_B^{-1} \colon M_{n\times 1}(\mathbb{R}) \to V$ defined by

$$\tau_B^{-1}\left(\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}\right) = c_1 b_1 \ldots + c_n b_n.$$

We have solved the issue of identifying V with column vectors as explicitly as possible.

2 Matrices and linear maps

There is another common identification in linear algebra: that of matrices and linear maps. The choice of basis also matters a great deal when defining this correspondence.

Let U be an m-dimensional real vector space with a basis B' and let $f\colon V \to U$ be a linear map. What we want is a matrix $X \in M_{m\times n}(\mathbb{R})$ such that if we take some $v \in V$, apply f to it, and interpret the result as a column vector in U it would be the same as first interpreting v as a column vector in V and then multiplying it by X. More concisely, we want the following diagram to commute.

$$\begin{tikzcd}
M_{n\times 1}(\mathbb{R}) \arrow[r, "L_X"] & M_{m\times 1}(\mathbb{R}) \\
V \arrow[u, "\tau_B"] \arrow[r, "f"'] & U \arrow[u, "\tau_{B'}"']
\end{tikzcd}$$

Here $L_X\colon M_{n\times 1}(\mathbb{R}) \to M_{n\times 1}(\mathbb{R})$ is just the result of multiplying X on the left, i.e. $L_X(C) = XC$ for all column vectors C. Whenever a diagram of this kind commutes we say that X represents f with respect to the bases B, B'. In the special case where $U = V$ and $B = B'$ then we can just say that X represents f in the basis B.

3 Change of basis

A natural question arises. Suppose we have some vector $v \in V$ expressed in the basis B. How do we express it in B', another basis of V? More precisely, given $\tau_B(v)$ we want to get $\tau_{B'}(v)$, ideally using an $n \times n$ matrix P. This is called a *change-of basis matrix* and it makes the following diagram commute.

$$\begin{tikzcd}
M_{n\times 1}(\mathbb{R}) \arrow[r, "L_P"] & M_{n\times 1}(\mathbb{R}) \\
V \arrow[u, "\tau_B"] \arrow[ur, "\tau_{B'}"']
\end{tikzcd}$$

This matrix P always exists; if $B' = \{b'_1, \dots, b'_n\}$ then setting the i-th column of P to be $\tau_B(b'_i)$ works. For our purposes, it is not relevant to know how we construct P, it suffices to know that it exists and satisfies the commutativity of the diagram. (Adhering to the philosophy of category theory, we do not care about what the object is, only its relations to other objects).

The property that P satisfies is that $\tau_{B'} = L_P \circ \tau_B$, call this equation (1). It follows that if we want to do the inverse process, change the basis from B' to B, then we would need a matrix P^* that satisfies $\tau_B = L_{P^*} \circ \tau_{B'}$, call this equation (2). By substituting equation (1) into equation (2) we get $\tau_B = L_{P*} \circ L_P \circ \tau_B$. Since τ_B is invertible, and since $L_{P*} \circ L_P = L_{P^*P}$, we have that L_{P^*P} must be an identity function. Similarly, if one substitutes equation (2) into equation (1) we get that L_{PP^*} is an identity. These both implies that P is invertible and that $P^* = P^{-1}$, as one would expect.

We have solved the problem of "translating" from a basis to another in the context of vectors. An analogous problem arises with linear maps. Let $f\colon V \to V$ be a linear map and let X be an $n \times n$ real matrix such that X represents f in the basis B. How do we represent f in B', another basis of V? We want to find another square matrix Y such that the following diagram commutes.

$$\begin{tikzcd}
M_{n\times 1}(\mathbb{R}) \arrow[r, "L_Y"] & M_{n\times 1}(\mathbb{R}) \\
V \arrow[u, "\tau_{B'}"] \arrow[r, "f"'] & V \arrow[u, "\tau_{B'}"']
\end{tikzcd}$$

To do this we can use the change-of-basis matrix as well. Start with the diagram defining X.

$$\begin{tikzcd}
M_{n\times 1}(\mathbb{R}) \arrow[r, "L_X"] & M_{n\times 1}(\mathbb{R}) \\
V \arrow[u, "\tau_B"] \arrow[r, "f"'] & V \arrow[u, "\tau_B"']
\end{tikzcd}$$

Let P be the change-of-basis matrix from B to B'. Then, by definition of P, the following diagram must commute as well.

$$\begin{tikzcd}
M_{n\times 1}(\mathbb{R}) \arrow[r, "L_X"] & M_{n\times 1}(\mathbb{R}) \arrow[r, "L_P"] & M_{n\times 1}(\mathbb{R}) \\
V \arrow[u, "\tau_B"] \arrow[r, "f"'] & V \arrow[u, "\tau_B"'] \arrow[ur, "\tau_{B'}"'] &
\end{tikzcd}$$

Similarly, we can bring P^{-1} into play as follows.

$$\begin{tikzcd}
M_{n\times 1}(\mathbb{R}) \arrow[r, "L_{P^{-1}}"] & M_{n\times 1}(\mathbb{R}) \arrow[r, "L_X"] & M_{n\times 1}(\mathbb{R}) \arrow[r, "L_P"] & M_{n\times 1}(\mathbb{R}) \\
& V \arrow[ul, "\tau_{B'}"] \arrow[u, "\tau_B"] \arrow[r, "f"'] & V \arrow[u, "\tau_B"'] \arrow[ur, "\tau_{B'}"'] &
\end{tikzcd}$$

It is straightforward to verify that this whole diagram commutes. We can take only the outer arrows to conclude that the matrix Y we were looking for is just $P^{-1}XP$. We have proved the following.

Theorem 3.1. *Let f be an endomorphism of V. If X is a matrix representing f in the basis B then for any basis B' of V there exists an invertible matrix P such that $P^{-1}XP$ represents f in the basis B'.*

4 The center of $GL_n\mathbb{R}$

Suppose we want to find all the matrices $Z \in GL_n(\mathbb{R})$ such that, for all $P \in GL_n(\mathbb{R})$ we have that Z and P commute, i.e. $ZP = PZ$. This condition can equivalently stated as $Z = P^{-1}ZP$. We say that Z is in the *center* of $GL_n(\mathbb{R})$. What does this mean in terms of linear maps?

Here is the answer. We consider V to be $\mathbb{R}^n$ and hence B is now an arbitrary basis of $\mathbb{R}^n$. Define g to be an endomorphism of $\mathbb{R}^n$ such that Z represents g in the basis B; this is done by letting $g \coloneqq \tau_B^{-1} \circ L_Z \circ \tau_B$. Then, for any basis B' of $\mathbb{R}^n$ there is an invertible matrix P such that $P^{-1}ZP$ represents g in the basis B'. But then if $Z = P^{-1}ZP$ this means that Z represents g in the basis B' as well. So g is an endomorphism with the property that there is a single matrix Z that represents it in every basis of $\mathbb{R}^n$(!). Equivalently, Z has the property of always representing the same endomorphism, no matter the basis.

We take advantage of this fact in the following way. Recall that $S \coloneqq \{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{R}^n$. Define, for all integers k with $1 \le k \le n$,

$$S_k \coloneqq \{e_1, \dots, -e_k, \dots e_n\}.$$

So S_k is the result of negating e_k in the standard basis; this still results in a valid basis of $\mathbb{R}^n$.

Fix some $1 \le k \le n$. We know that Z represents g in both S and S_k. Hence the identities $L_Z \circ \tau_S = \tau_S \circ g$ and $L_Z \circ \tau_{S_k} = \tau_{S_k} \circ g$ are both true. Solving for g in both equations we get that

$$g = \tau_S^{-1} \circ L_Z \circ \tau_S = \tau_{S_k}^{-1} \circ L_Z \circ \tau_{S_k}.$$

Suppose we input e_k to this function, so that we have the following chain of reasoning.

$$\tau_S^{-1} \circ L_Z \circ \tau_S(e_k) = \tau_{S_k}^{-1} \circ L_Z \circ \tau_{S_k}(-(-e_k))$$

$$\tau_S^{-1} \circ L_Z \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \tau_{S_k}^{-1} \circ L_Z \begin{bmatrix} 0 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix}.$$

If $(Z)_{ij} = z_{ij}$ for all $1 \leq i, j \leq n$.

$$\tau_S^{-1} \begin{bmatrix} z_{1k} \\ \vdots \\ z_{nk} \end{bmatrix} = \tau_{S_k}^{-1} \begin{bmatrix} -z_{1k} \\ \vdots \\ -z_{nk} \end{bmatrix}$$

$$z_{1k}e_1 + \ldots + \cancel{z_{kk}e_k} + \ldots + z_{nk}e_n = -z_{1k}e_1 - \ldots \cancel{-z_{kk}(-e_k)} - \ldots - z_{nk}e_n$$
$$2z_{1k}e_1 + \ldots + 2z_{(k-1)k}e_{k-1} + 2z_{(k+1)k}e_{k+1} + \ldots + 2z_{nk}e_n = 0$$

The standard basis without e_k still remains a linearly independent set. Hence all the scalars in the equation above are equal to zero. Repeating this reasoning for all $1 \leq k \leq n$ tells us that all the z's are zero, except those of the form z_{kk}, i.e. the matrix Z is diagonal.

We can use a similar trick to get more information about Z. Define, for integers k, l such that $1 \leq k < l \leq n$,

$$S_{kl} := \{e_1, \ldots, e_l, \ldots, e_k, \ldots, e_n\}.$$

So S_{kl} is the result of swapping e_k and e_l. Again, Z represents g in both S and S_{kl}. So we can say that

$$\tau_S^{-1} \circ L_Z \circ \tau_S = \tau_{S_{kl}}^{-1} \circ L_Z \circ \tau_{S_{kl}}.$$

Inputting e_k to this function yields $z_{kk}e_k = z_{ll}e_k$. This implies $z_{kk} = z_{ll}$ for all integers k, l with $1 \leq k < l \leq n$. So Z is not only diagonal but all of its diagonal entries are equal. As $Z \in GL_n(\mathbb{R})$, we conclude that Z must be a nonzero multiple of the identity. And indeed, all multiples of the identity commute with all matrices. Hence we have completely characterized the *center* of $GL_n(\mathbb{R})$.