# Change of basis with commutative diagrams

Hernán Ibarra Mejia

July 12, 2022

## 0   What are commutative diagrams?

A commutative diagram is just a helpful way to state that some functions (or rather, compositions of functions) are the same. Many of the concepts you have encountered (and ones that you will encounter) have an elegant representation in terms of commutative diagrams.

Suppose we have some sets and some functions between them. I could list all the sets and then state the names of functions along with the information of their domain and codomain. Much more economically, I could just show the following diagram.

$$
\begin{array}{ccccccccc}
A & \xrightarrow{\;f\;} & B & \xrightarrow{\;g\;} & C & \xrightarrow{\;h\;} & D & \xrightarrow{\;j\;} & E \\
\downarrow{\scriptstyle l} & & \downarrow{\scriptstyle m} & & \downarrow{\scriptstyle n} & & \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle q} \\
A' & \xrightarrow{\;r\;} & B' & \xrightarrow{\;s\;} & C' & \xrightarrow{\;t\;} & D' & \xrightarrow{\;u\;} & E'
\end{array}
$$

Not only do you immediately understand what are the sets in question, you can also deduce things like $h$ is a function from $C$ to $D$.

If we say that the diagram above commutes, we mean that "paths" that have the same starting point and endpoint are equal. For instance, we see that both the composition $n \circ g \circ f$ and $s \circ r \circ l$ take (elements of) $A$ to (elements of) $C'$; hence the commutativity of the diagram says that $n \circ g \circ f = s \circ r \circ l$. We can derive many equalities of this kind, another example would be $p \circ h = t \circ n$. It is certainly easier to draw a diagram and say it is commutative than listing all sets, with the functions connecting them, with the equalities of compositions that hold.

*Exercise* 1. Suppose $G$ and $H$ are sets, such that we have some binary operations $m_G \colon G \times G \to G$ and $m_H \colon H \times H \to H$ that make $G$ and $H$ into groups. We let $\varphi \colon G \to H$ be a function, and define $\varphi \times \varphi \colon G \times G \to H \times H$ by the formula

$$\varphi \times \varphi\,(g_1, g_2) := (\varphi(g_1), \varphi(g_2)), \text{ for all } (g_1, g_2) \in G \times G.$$

Prove that $\varphi$ is a group homomorphism if and only if the following diagram commutes.

$$
\begin{array}{ccc}
G \times G & \xrightarrow{\varphi \times \varphi} & H \times H \\
{\scriptstyle m_G}\downarrow & & \downarrow{\scriptstyle m_H} \\
G & \xrightarrow{\quad\varphi\quad} & H
\end{array}
$$

# 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 (if you don't know what that means, take $V = \mathbb{R}^n$); and we suppose $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ 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 $c_1, \ldots c_n \in \mathbb{R}$ such that $v = c_1 \mathbf{b}_1 + \ldots + c_n \mathbf{b}_n$. Then we let

$$
\tau_B(v) = \tau_B(c_1 \mathbf{b}_1 + \ldots + c_n \mathbf{b}_n) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.
$$

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

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

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

It is worth pointing out that $\tau_B$ is a genuine bijection since it has an inverse $\tau_B^{-1} \colon M_{n \times 1}(\mathbb{R}) \to V$ defined by
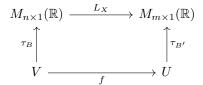
$$
\tau_B^{-1}\left( \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \right) = c_1 b_1 \ldots + c_n b_n.
$$

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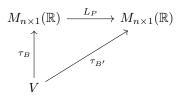
Let $U$ be an $m$-dimensional real vector space with a basis $B'$ and let $f\colon V \to U$ be a linear map. What we want is a matrix $X \in M_{m\times n}(\mathbb{R})$ such that if we take some $v \in V$, apply $f$ to it, and interpret the result as a column vector in $U$ it would be the same as first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

$$
\begin{array}{ccc}
M_{n\times 1}(\mathbb{R}) & \xrightarrow{\;L_X\;} & M_{m\times 1}(\mathbb{R}) \\
\tau_B \uparrow & & \uparrow \tau_{B'} \\
V & \xrightarrow[\;f\;]{} & U
\end{array}
$$

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.

$$
\begin{array}{ccc}
M_{n\times 1}(\mathbb{R}) & \xrightarrow{\;L_P\;} & M_{n\times 1}(\mathbb{R}) \\
\tau_B \uparrow & \nearrow \tau_{B'} & \\
V & &
\end{array}
$$

This matrix $P$ always exists; if $B' = \{b'_1, \ldots, b'_n\}$ 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 $\tau_B$ 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.

$$
\begin{array}{ccc}
M_{n\times 1}(\mathbb{R}) & \xrightarrow{\ L_Y\ } & M_{n\times 1}(\mathbb{R}) \\
\uparrow{\scriptstyle \tau_{B'}} & & \uparrow{\scriptstyle \tau_{B'}} \\
V & \xrightarrow{\ f\ } & V
\end{array}
$$

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

$$
\begin{array}{ccc}
M_{n\times 1}(\mathbb{R}) & \xrightarrow{\ L_X\ } & M_{n\times 1}(\mathbb{R}) \\
\uparrow{\scriptstyle \tau_{B}} & & \uparrow{\scriptstyle \tau_{B}} \\
V & \xrightarrow{\ f\ } & V
\end{array}
$$

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

$$
\begin{array}{ccccc}
M_{n\times 1}(\mathbb{R}) & \xrightarrow{\ L_X\ } & M_{n\times 1}(\mathbb{R}) & \xrightarrow{\ L_P\ } & M_{n\times 1}(\mathbb{R}) \\
\uparrow{\scriptstyle \tau_{B}} & & \uparrow{\scriptstyle \tau_{B}} & \nearrow{\scriptstyle \tau_{B'}} & \\
V & \xrightarrow{\ f\ } & V & &
\end{array}
$$

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

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 $g := \tau_B^{-1} \circ L_Z \circ \tau_B$. 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 $S := \{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{R}^n$. Define, for all integers $k$ with $1 \le k \le n$,

$$S_k := \{e_1, \ldots, -e_k, \ldots e_n\}.$$

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 $GL_n(\mathbb{R})$.