# Matrices that represent the same Linear Transformation and diagonalization

— Joaquín Gómez —

— February 24, 2025 —

Let $T : V \to W$ be a **linear transformation** from $V \subseteq \mathbb{R}^n$ to $W \subseteq \mathbb{R}^m$, and let $(e_1, ..., e_n)$ and $(w_1, ..., w_m)$ be the bases of $V$ and $W$ respectively.

Now, let $A = [a_{ij}]_{i,j=1}^{n,m}$ be the **matrix representation** of $T$. If we apply the transformation on the $k^{th}$ basis vector of $V$ we get

$$T(e_k) = \sum_{i=1}^{n} a_{ik} e_i \tag{1}$$

Now, suppose we form another basis using the basis of $V$, let this basis be $U = (u_1, ..., u_n)$ where each $u_k$ is given by

$$u_k = \sum_{j=1}^{n} c_{jk} e_j \tag{2}$$

Then $C = [c_{ij}]_{i,j=1}^{n}$ is the matrix for the change of basis, such that $U = EC$. Then, applying the transformation on $u_k$ we find

$$T(u_k) = T\left(\sum_{j=1}^{n} c_{jk} e_j\right) = \sum_{j=1}^{n} c_{jk} T(e_j) = \sum_{j=1}^{n} c_{jk} \sum_{i=1}^{n} a_{ij} e_i$$

reordering terms, we get

$$T(u_k) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{jk} a_{ij} e_i \tag{3}$$

We can set $U' = (T(u_1), ..., T(u_n))$, and $E' = (T(e_1), ..., T(e_n))$. If we look at $\sum_{j=1}^{n} c_{jk} T(e_j)$, we see that

$$U' = E'C, \quad E' = EA \longrightarrow U' = EAC$$

We have proved this above, this is the same as in (3).

Now, we define another matrix $B = [b_{ij}]_{i,j=1}^{n,m}$ that represents the transformation with respect to the basis of $U$, this is, $B$ is the equivalent of $A$ in $U$.

$$T(u_k) = \sum_{j=1}^{n} b_{jk} u_j \tag{4}$$

In the matrix form, this is

$$U' = UB$$

so, going back to our equations, and recalling that $U = EC \rightarrow E = UC^{-1}$

$$U' = EAC = UC^{-1}AC \rightarrow UB = UC^{-1}AC$$

Then

$$B = C^{-1}AC \tag{5}$$

This is the relationship between the matrix representation of $T$ in $U$ ($B$) and the matrix representation of $T$ in $V$ ($A$).

## Application with eigenvectors

Let $M$ be a **symmetric** matrix, then

$$Mx = \lambda x, \qquad My = \mu y$$

Now, consider the following dot product

$$(Mx, My) = (\lambda x, My) = \lambda(x, My) = (Mx, \mu y) = \mu(Mx, y)$$

Then

$$\lambda(x^T My) = \mu(x^T M^T y)$$

But $x^T y = y^T x$, and $M^T = M$, **because $M$ is symmetric** (this is the importance of the constraint). So we rewrite the equation by transposing the right-hand side, we obtain

$$\lambda(x^T My) = \mu(x^T My) \longrightarrow x^T My(\lambda - \mu)$$

Our constraint is that $\mu \neq \lambda$, this means that $x \cdot y = 0$, in other words, both eigenvectors are perpendicular.

This initial proof is important because it sets the bases for the next theorem.

**Theorem:** any matrix $M$ which is symmetric can be diagonalized by a diagonal matrix of eigenvalues and an orthogonal matrix of eigenvectors.

To prove this we use the fact that $M$ has orthogonal eigenvectors, this means that we have a basis of eigenvector $U = (u_1, ..., u_n)$. Suppose that $T$ is a transformation on this basis, we know that the eigenvectors $u_1, ..., u_n$ are obtained by

$$T(u) = \lambda u$$

In the case of matrices, $T(u) = Mu$, but this notation implies that the summation given in (4) has only one term in this situation. We get that the matrix that applies the same transformation in the basis $U$ (which in the previous demonstration was $B$), is a matrix consisting of the eigenvalues in the main diagonal.

$$T(u_k) = \sum_{j=1}^{n} b_{jk} u_j = \lambda_k u_k \tag{6}$$

So, collecting all the $k$ terms in (6), we get

$$U' = \Lambda U$$

Then, given that our linear transformation $T$ is represented by a matrix $M$ in the basis $E$

$$E' = EM$$

Hence each eigenvector is given by a linear combination of elements in $E$

$$u_k = \sum_{j=1}^{n} s_{jk} e_j \tag{7}$$

We call the matrix of eigenvector $S$, which has each eigenvector in its columns as $S = [u_1 ... u_n]$. Such that

$$U = ES, \qquad E = US^{-1}$$

because $T(u_k) = \sum_{j=1}^{n} s_{jk} T(e_j)$ for each eigenvector.

Which finally yields for

$$U' = EMS \longrightarrow \Lambda U = US^{-1}MS$$

This proves the theorem.

Note that

$$\Lambda = S^{-1}MS, \qquad M = S\Lambda S^{-1}$$

And, because $M$ is symmetric, then $S$ is orthogonal. So $S^{-1} = S^T$ and

$$\Lambda = S^T MS, \qquad M = S\Lambda S^T \tag{8}$$

This simplifies the calculations a lot, this kind of matrix is a good-behaved one. Also, if the matrix is complex, this requires $M$ to be *hermitian*. The hermitian property satisfies an operator

$$M^* = M$$

This operator takes the transpose but also the conjugate, the conjugate is element-wise. Conjugation is an operation on a complex number such that

$$c = a + ib, \quad \bar{c} = a - ib$$

This is essentially the “reflection” of the complex number. If you conjugate a matrix $A$, by doing $\bar{A}$, you will get all of its elements conjugated. Then taking the transpose must yield $\bar{A}^T = A$ if it is hermitian. If $A$ its hermitian and represents a linear transformation $T$, then $T$ is hermitian (we are not going into more detail here).

The point is that if $M$ is hermitian, then the equations work too and become

$$\Lambda = S^* M S, \qquad M = S \Lambda S^*$$