MATRIX REPRESENTATION OF LINEAR OPERATORS: CHANGE OF BASIS

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

References: edX online course MIT 8.05.1x Week 3.

Sheldon Axler (2015), *Linear Algebra Done Right*, 3rd edition, Springer. Chapter 3.

We've seen that the matrix representation of a linear operator depends on the basis we've chosen within a vector space V. We now look at how the matrix representation changes if we change the basis. In what follows, we'll consider two sets of basis vectors $\{v\}$ and $\{u\}$ and two operators A and B. Operator A transforms the basis $\{v\}$ into the basis $\{u\}$, while B does the reverse. That is

$$Av_i = u_i \tag{0.1}$$

$$Bu_i = v_i \tag{0.2}$$

for all $i = 1, \ldots, n$. From this definition, we can see that $A = B^{-1}$ and $B = A^{-1}$, since

$$u_i = Av_i = ABu_i \tag{0.3}$$

$$v_i = Bu_i = BAv_i \tag{0.4}$$

Theorem 1. *An operator (like A or B above) that transforms one set of basis vectors into another has the same matrix representation in both bases.*

Proof. In matrix form, we have (remember we're using the summation convention on repeated indices):

$$Av_i = A_{ji}(\{v\})\, v_j \tag{0.5}$$

$$Au_i = A_{ji}(\{u\})\, u_j \tag{0.6}$$

Note that the matrix elements depend on *different* bases in the two equations.

We can now operate with A again, using 0.1, to get

$$Au_i = A(Av_i) \tag{0.7}$$
$$= A\left(A_{ji}(\{v\})v_j\right) \tag{0.8}$$
$$= A_{ji}(\{v\})Av_j \tag{0.9}$$
$$= A_{ji}(\{v\})u_j \tag{0.10}$$

Comparing the last line with 0.6, we see that

$$A_{ji}(\{v\}) = A_{ji}(\{u\})$$

Since the matrix elements are just numbers, this means that the elements in the two matrices $A_{ji}(\{v\})$ and $A_{ji}(\{u\})$ are the same.

We could do the same analysis using the B operator with the same result:

$$B_{ji}(\{v\}) = B_{ji}(\{u\}) \tag{0.11}$$

□

We can now turn to the matrix representations of a general operator T in two different bases. In this case, T can perform any linear transformation, so it doesn't necessarily transform one set of basis vectors into another set of basis vectors. Consider first the case where T operates on each set of basis vectors given above:

$$Tv_i = T_{ji}(\{v\})v_j \tag{0.12}$$
$$Tu_i = T_{ji}(\{u\})u_j \tag{0.13}$$

Unless T is an operator like A or B above, in general $T_{ji}(\{v\}) \neq T_{ji}(\{u\})$. We can see how these two matrices are related by using operators A and B above to write

$$Tu_i = T\left(A_{ji}v_j\right) \tag{0.14}$$
$$= A_{ji}Tv_j \tag{0.15}$$
$$= A_{ji}T_{kj}(\{v\})v_k \tag{0.16}$$
$$= A_{ji}T_{kj}(\{v\})Bu_k \tag{0.17}$$
$$= A_{ji}T_{kj}(\{v\})A^{-1}u_k \tag{0.18}$$
$$= A_{ji}T_{kj}(\{v\})A^{-1}_{pk}u_p \tag{0.19}$$
$$= \left[A^{-1}_{pk}T_{kj}(\{v\})A_{ji}\right]u_p \tag{0.20}$$
$$= T_{pi}(\{u\})u_p \tag{0.21}$$

We don't need to specify the basis for the A or B matrices since the matrices are the same in both bases as we just saw above. The last line is just the expansion of Tu_i in terms of the $\{u\}$ basis. In the penultimate line, we see that the quantity in square brackets is the product of 3 matrices:

$$A^{-1}_{pk}T_{kj}(\{v\})A_{ji} = \left[A^{-1}T(\{v\})A\right]_{pi} \tag{0.22}$$

The required transformation is therefore

$$T(\{u\}) = A^{-1}T(\{v\})A \tag{0.23}$$

where $u_i = Av_i$.

As a check, note that if $T = A$ or $T = B = A^{-1}$, we reclaim the result in the theorem above, namely that $A(\{u\}) = A(\{v\})$ and $B(\{u\}) = B(\{v\})$.

Trace and determinant. The trace of a matrix is the sum of its diagonal elements, written as $\mathrm{tr}\, T$. A useful property of the trace is that

$$\mathrm{tr}\ (AB) = \mathrm{tr}\ (BA) \tag{0.24}$$

We can prove this by looking at the components. If $C = AB$ then

$$C_{ij} = A_{ik}B_{kj} \tag{0.25}$$

The trace of C is the sum of its diagonal elements, written as C_{ii}, so

$$\begin{aligned}
\mathrm{tr}\, C &= \mathrm{tr}\ (AB) && (0.26)\\
&= A_{ik}B_{ki} && (0.27)\\
&= B_{ki}A_{ik} && (0.28)\\
&= [BA]_{kk} && (0.29)\\
&= \mathrm{tr}\ (BA) && (0.30)
\end{aligned}$$

From this we can generalize to the case of the trace of a product of any number of matrices and obtain the cyclic rule:

$$\mathrm{tr}\,(A_1A_2\ldots A_n) = \mathrm{tr}\,(A_nA_1A_2\ldots A_{n-1}) \tag{0.31}$$

Going back to 0.23, we have

$$\operatorname{tr} T(\{u\}) = \operatorname{tr}\left(A^{-1}T(\{v\})A\right) \tag{0.32}$$
$$= \operatorname{tr}\left(AA^{-1}T(\{v\})\right) \tag{0.33}$$
$$= \operatorname{tr} T(\{v\}) \tag{0.34}$$

Thus the trace of any linear operator is invariant under a change of basis.

For the determinant, we have the results that the determinant of a product of matrices is equal to the product of the determinants, and the determinant of a matrix inverse is the reciprocal of the determinant of the original matrix. Therefore

$$\det(T(\{u\})) = \det\left(A^{-1}T(\{v\})A\right) \tag{0.35}$$
$$= \frac{\det A}{\det A}\det T(\{v\}) \tag{0.36}$$
$$= \det T(\{v\}) \tag{0.37}$$

Thus the determinant is also invariant under a change of basis.

PINGBACKS

Pingback: Projection operators

Pingback: Unitary operators

Pingback: Diagonalization of matrices

Pingback: Unitary operators: active and passive transformations of an operator

Pingback: Pauli matrices: properties