13.5 Matrix Representation of Operators

To get an extra handle on linear operators, we sometimes represent them as matrices. This is actually a completely general statement: after choosing a basis on the vector space V, *any* linear operator $T \in \mathcal{L}(V)$ can be represented by a particular matrix. This representation carries *all* the information about the linear operator. The matrix form of the operator can be very useful for explicit computations. The only downside of matrix representations is that they depend on the chosen basis. On the upside, a clever choice of basis may result in a matrix representation of unusual simplicity, which can be quite valuable. Additionally, some quantities computed easily from the matrix representation of an operator do not depend on the choice of basis.

The **matrix representation** of a linear operator $T \in \mathcal{L}(V)$ is a matrix whose components $T_{ij}(\{v\})$ are read from the action of the operator T on each of the elements of a basis $(v_1, \ldots, v_n)$ of V. The notation $T_{ij}(\{v\})$ reflects the fact that the matrix components depend on the choice of basis. If the choice of basis is clear by the context, we simply write the matrix components as T_{ij}. The rule that defines the matrix is simple:

> Rule: The jth column of the matrix T is the list of components of Tv_j when expanded along the basis.

$$\boxed{\begin{pmatrix} \cdots & \cdots & T_{1j} & \cdots \\ \cdots & \cdots & T_{2j} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & T_{nj} & \vdots \end{pmatrix}, \qquad Tv_j = T_{1j}\, v_1 + T_{2j}\, v_2 + \cdots + T_{nj}\, v_n.} \tag{13.5.1}$$

Example 13.17. *Matrix representation for an operator in* $\mathbb{R}^3$.
The action of T on some basis vectors (v_1, v_2, v_3) of $\mathbb{R}^3$ is given by

$$\begin{aligned} Tv_1 &= -v_1 + 7v_3, \\ Tv_2 &= 2v_1 + v_2 + 3v_3, \\ Tv_3 &= 6v_1 - 5v_2 + 8v_3. \end{aligned} \tag{13.5.2}$$

The matrix representation of T is then

$$\begin{pmatrix} -1 & 2 & 6 \\ 0 & 1 & -5 \\ 7 & 3 & 8 \end{pmatrix}. \tag{13.5.3}$$

This follows by direct application of the rule. From the action of T on v_1, for example, we have

$$Tv_1 = -v_1 + 7v_3 = -v_1 + 0 \cdot v_2 + 7v_3 = T_{11}v_1 + T_{21}v_2 + T_{31}v_3, \tag{13.5.4}$$

allowing us to read the first column of the matrix. The other columns follow similarly.

□

The equation in (13.5.1) can be written more briefly as

$$\boxed{Tv_j = \sum_{i=1}^{n} T_{ij}\, v_i.} \tag{13.5.5}$$

We write the summation sign explicitly for clarity; it is sometimes omitted on account of the summation convention that states that repeated indices are understood to be summed over. While operators are represented by matrices, vectors in V are represented by **column vectors**: the entries on the column vector are the components of the vector along the basis vectors. For a vector $a \in V$,

$$a = a_1 v_1 + \cdots + a_n v_n \quad \longleftrightarrow \quad a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}. \tag{13.5.6}$$

It is a simple consequence that the basis vector v_k is represented by a column vector of zeroes, with a one on the kth entry:

$$v_k = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k\text{th}. \tag{13.5.7}$$

With this you can now see why our key definition $Tv_j = \sum_{i=1}^{n} T_{ij} v_i$ is consistent with the familiar rule for multiplication of a matrix times a vector:

$$Tv_j = \begin{pmatrix} T_{11} & \cdots & T_{1j} & \cdots & T_{1n} \\ T_{21} & \cdots & T_{2j} & \cdots & T_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ T_{n1} & \cdots & T_{nj} & \cdots & T_{nn} \end{pmatrix} \begin{array}{c} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \\ j\text{th} \\ \\ \\ \end{array} = \begin{pmatrix} T_{1j} \\ T_{2j} \\ \vdots \\ T_{nj} \end{pmatrix} \tag{13.5.8}$$

$$= T_{1j} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + T_{2j} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + T_{nj} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = T_{1j}v_1 + T_{2j}v_2 + \ldots + T_{nj}v_n.$$

Exercise 13.7. *Verify that in any basis the matrix representation of the identity operator is a diagonal matrix with an entry of one at each element of the diagonal and zero elsewhere.*

The rules for representations are not only consistent with the formula for multiplication of matrices times vectors; they actually *imply* this familiar formula. In fact, they also imply the famous rule for matrix multiplication. We discuss both of these now.

Consider vectors a, b that, expanded along the basis $(v_1, \ldots, v_n)$, read

$$\begin{aligned} a &= a_1v_1 + \cdots + a_nv_n, \\ b &= b_1v_1 + \cdots + b_nv_n. \end{aligned} \tag{13.5.9}$$

Assume the vectors are related by the equation

$$b = Ta. \tag{13.5.10}$$

We want to see how this looks in terms of the representations of T, a, and b. We have

$$b = Ta = T\sum_j a_jv_j = \sum_j a_j\, Tv_j = \sum_{i,j} a_j\, T_{ij}v_i = \sum_i \Big(\sum_j T_{ij}a_j\Big)v_i. \tag{13.5.11}$$

The object in parentheses is the ith component of b:

$$b_i = \sum_j T_{ij}a_j. \tag{13.5.12}$$

This is how $b = Ta$ is represented; we see on the right-hand side the familiar product of the matrix for T and the column vector for a.

Let us now examine the product of two operators and their matrix representation. Consider the operator TS acting on v_j:

$$(TS)v_j = T(Sv_j) = T\sum_p S_{pj}v_p = \sum_p S_{pj}\,Tv_p = \sum_p S_{pj}\sum_i T_{ip}v_i \tag{13.5.13}$$

so that changing the order of the sums we find that

$$(TS)v_j = \sum_i\Big(\sum_p T_{ip}S_{pj}\Big)\,v_i. \tag{13.5.14}$$

Using the identification implicit in (13.5.5), we see that the object in parentheses is $(TS)_{ij}$, the i, j element of the matrix that represents TS. Therefore, we find that

$$(TS)_{ij} = \sum_p T_{ip}S_{pj}, \tag{13.5.15}$$

which is precisely the familiar formula for matrix multiplication. The matrix that represents TS is the product of the matrix that represents T and the matrix that represents S, in that order.

Changing basis and its effect on matrix representations While matrix representations are very useful for concrete visualization, they are basis dependent. It is a good idea to try to determine if there are quantities that can be calculated using a matrix representation that are, nevertheless, guaranteed to be basis independent. One such quantity is the **trace** of the matrix representation of a linear operator. The trace is the sum of the matrix elements on the diagonal. Remarkably, that sum is the same independent of the basis used. This allows us to speak of the trace of an *operator*. The **determinant** of a matrix representation is also basis independent. We can therefore speak of the determinant of an operator. We will prove the basis independence of the trace and the determinant once we learn how to relate matrix representations in different bases.

Let us then consider the effect of a change of basis on the matrix representation of an operator. Consider a vector space V and two sets of basis vectors: $(v_1, \ldots, v_n)$ and $(u_1, \ldots, u_n)$. Consider then two linear operators $A, B \in \mathcal{L}(V)$ such that for any $i = 1, \ldots, n$, A acting on v_i gives u_i, and B acting on u_i gives v_i:

$$A: \begin{array}{ccc} v_1 & \ldots & v_n \\ \downarrow & \ldots & \downarrow \\ u_1 & \ldots & u_n \end{array}, \qquad B: \begin{array}{ccc} v_1 & \ldots & v_n \\ \uparrow & \ldots & \uparrow \\ u_1 & \ldots & u_n \end{array}. \tag{13.5.16}$$

This can also be written as

$$Av_k = u_k, \qquad Bu_k = v_k, \quad \text{for all } k \in \{1, \ldots, n\}. \tag{13.5.17}$$

These relations define the operators A and B completely: we have stated how they act on basis sets. We now verify the obvious: A and B are inverses of each other. Indeed,

$$\begin{aligned} BAv_k &= B(Av_k) = Bu_k \ = v_k, \\ ABu_k &= A(Bu_k) = Av_k \ = u_k, \end{aligned} \tag{13.5.18}$$

being valid for all k, shows that

$$BA = \mathbb{1} \quad \text{and} \quad AB = \mathbb{1}. \tag{13.5.19}$$

Thus, B is the inverse of A, and A is the inverse of B.

Operators like A or B that map one basis into another, vector by vector, have a remarkable property: *their matrix representations are the same in each of the bases they relate*. Let us prove this for A. By definition of matrix representations, we have

$$A\, v_k \ = \ \sum_i A_{ik}(\{v\}) v_i, \qquad \text{and} \qquad A\, u_k \ = \ \sum_i A_{ik}(\{u\}) u_i. \tag{13.5.20}$$

Since $u_k = Av_k$, we then have, acting with another A,

$$Au_k = A(Av_k) = A \sum_i A_{ik}(\{v\}) v_i = \sum_i A_{ik}(\{v\})\, Av_i = \sum_i A_{ik}(\{v\})\, u_i. \tag{13.5.21}$$

Comparison with the second equation immediately above yields the claimed

$$A_{ik}(\{u\}) = A_{ik}(\{v\}). \tag{13.5.22}$$

The same holds for the B operator. We can simply call A_{ij} and B_{ij} the matrices that represent A and B because these matrices are the same in the $\{v\}$ and $\{u\}$ bases, and these are the only bases at play here. On account of (13.5.19), these are matrix inverses:

$$B_{ij}A_{jk} = \delta_{ik} \quad \text{and} \quad A_{ij}B_{jk} = \delta_{ik}, \tag{13.5.23}$$

and we can write $B_{ij} = (A^{-1})_{ij}$. At this point we will use the convention that repeated indices are summed over to avoid clutter.

We can now apply these preparatory results to the matrix representations of the operator T. We have, by definition,

$$Tv_k = T_{ik}(\{v\})\, v_i. \tag{13.5.24}$$

We need to calculate Tu_k in order to read the matrix representation of T on the u basis:

$$Tu_k = T_{ik}(\{u\})\, u_i. \tag{13.5.25}$$

Computing the left-hand side, using the linearity of the operator T, we have

$$Tu_k = T(Av_k) = T(A_{jk}v_j) = A_{jk}Tv_j = A_{jk}\, T_{pj}(\{v\})\, v_p. \tag{13.5.26}$$

We need to express the rightmost v_p in terms of u vectors. For this,

$$v_p = Bu_p = B_{ip}\, u_i = (A^{-1})_{ip}\, u_i \tag{13.5.27}$$

so that

$$Tu_k = A_{jk}\, T_{pj}(\{v\})\, (A^{-1})_{ip}\, u_i \;=\; (A^{-1})_{ip}\, T_{pj}(\{v\})\, A_{jk}\, u_i, \tag{13.5.28}$$

where we reordered the matrix elements to clarify the matrix products. All in all,

$$Tu_k = \big(A^{-1}T(\{v\})A\big)_{ik}\, u_i, \tag{13.5.29}$$

which, comparing with (13.5.25), allows us to read

$$T_{ij}(\{u\}) = \big(A^{-1}T(\{v\})A\big)_{ij}. \tag{13.5.30}$$

Omitting the indices, this matrix relation is written as

$$\boxed{T(\{u\}) \;=\; A^{-1}T(\{v\})\,A, \quad \text{when} \quad u_i = A\,v_i.} \tag{13.5.31}$$

Note that in the first relation A stands for a matrix, but in $u_i = Av_i$ it stands for an operator. This is the result we wanted to obtain. In general, if two matrices R, S are related by $S = M^{-1}RM$ for some matrix M, we say that S is obtained from R by a similarity transformation generated by M. In this language the matrix representation $T(\{u\})$ is obtained from the matrix representation $T(\{v\})$ by a similarity transformation generated by the matrix A that represents the operator that changes the basis from $\{v\}$ to $\{u\}$.

The trace of a matrix is equal to the sum of its diagonal entries. Thus, the trace of T is given by T_{ii}, with the sum over i understood. The trace is cyclic when acting on the product of various matrices:

$$\text{tr}(S_1S_2\ldots S_{k-1}S_k) = \text{tr}(S_kS_1S_2\ldots S_{k-1}). \tag{13.5.32}$$

This result is a simple consequence of $\text{tr}(S_1S_2) = \text{tr}(S_2S_1)$, which is easily verified by writing out the explicit products and taking the traces. With the help of (13.5.31) and the cyclicity of the trace, the basis independence of the trace follows quickly:

$$\text{tr}(T(\{u\})) \;=\; \text{tr}(A^{-1}T(\{v\})A) \;=\; \text{tr}(A\,A^{-1}T(\{v\})) \;=\; \text{tr}(T(\{v\})). \tag{13.5.33}$$

For the determinant we recall that

$$\det(S_1S_2) = (\det S_1)(\det S_2). \tag{13.5.34}$$

This means that $\det(S)\det(S^{-1}) = 1$ and that the determinant of the product of multiple matrices is also the product of determinants. From (13.5.31) we then find that

$$\det T(\{u\}) \;=\; \det(A^{-1}T(\{v\})A) = \det(A^{-1})\det T(\{v\})\det A \;=\; \det T(\{v\}), \tag{13.5.35}$$

showing that the determinant of the matrix that represents a linear operator is independent of the chosen basis.

Example 13.18. *Matrix representation for the harmonic oscillator $\hat{a}$ and $\hat{a}^\dagger$ operators.*
The one-dimensional simple harmonic oscillator state space is infinite-dimensional, and an *orthonormal* basis $\{e_1, e_2, \ldots\}$ is provided by the infinite set of nondegenerate energy eigenstates:

$$\{e_1, e_2, e_3, \ldots\} = \{\varphi_0, \varphi_1, \varphi_2, \ldots\}, \tag{13.5.36}$$

where, as explained in section 9.4,

$$\varphi_n = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n \varphi_0, \tag{13.5.37}$$

and φ_0 is the ground state wave function. Moreover, recalling the basic commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$, you can quickly recheck the result (9.4.26):

$$\begin{aligned} \hat{a}\,\varphi_n &= \sqrt{n}\,\varphi_{n-1}, \\ \hat{a}^\dagger \varphi_n &= \sqrt{n+1}\,\varphi_{n+1}. \end{aligned} \tag{13.5.38}$$

Given that $e_n = \varphi_{n-1}$ for $n = 1, \ldots$, the above relations become

$$\begin{aligned} \hat{a}\,e_n &= \sqrt{n-1}\,e_{n-1}, \\ \hat{a}^\dagger e_n &= \sqrt{n}\,e_{n+1}. \end{aligned} \tag{13.5.39}$$

The matrix representations now follow from the definition (13.5.5). Letting $\{\hat{a}\}_{mn}$ and $\{\hat{a}^\dagger\}_{mn}$ denote, respectively, the matrix elements of $\hat{a}$ and $\hat{a}^\dagger$, we have

$$\begin{aligned} \hat{a}\,e_n &= \sum_m \{\hat{a}\}_{mn} e_m = \sqrt{n-1}\,e_{n-1} \quad \Rightarrow \quad \{\hat{a}\}_{mn} = \sqrt{m}\,\delta_{m,n-1}, \\ \hat{a}^\dagger\, e_n &= \sum_m \{\hat{a}^\dagger\}_{mn}\, e_m = \sqrt{n}\, e_{n+1} \quad \Rightarrow \quad \{\hat{a}^\dagger\}_{mn} = \sqrt{n}\,\delta_{m,n+1}. \end{aligned} \tag{13.5.40}$$

The matrix $\{\hat{a}\}_{mn}$ is upper diagonal: the Kronecker delta $\delta_{m,n-1}$ tells us that for any row index m, the value of the column index n must be one unit higher. Similarly, the matrix $\{\hat{a}^\dagger\}_{mn}$ is lower diagonal. We have, explicitly,

$$\{\hat{a}\} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \quad \{\hat{a}^\dagger\} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{13.5.41}$$

The operators $\hat{a}$ and $\hat{a}^\dagger$ are Hermitian conjugates of each other. Their matrices are also Hermitian conjugates of each other: they are related by transposition and complex conjugation. In fact, being real, the matrices are just related by transposition. A Hermitian matrix is one left invariant by transposition and complex conjugation. As we will show in section 14.4, the matrix representation of a Hermitian operator is Hermitian when we use an orthonormal basis. The above representations imply that the number operator $\hat{N} = \hat{a}^\dagger\hat{a}$ is represented as the infinite diagonal matrix $\{\hat{N}\}$ = diag$(0, 1, 2, 3, \cdots)$. This is consistent with $\hat{N}\varphi_n = n\varphi_n$.

□

Exercise 13.8. *In example 13.16 and exercise 13.6, you found the right inverse for $\hat{a}$ and the left inverse for $\hat{a}^\dagger$, respectively. Write down the corresponding matrices as in (13.5.41).*