5C Upper-Triangular Matrices

In Chapter 3 we defined the matrix of a linear map from a finite-dimensional vector space to another finite-dimensional vector space. That matrix depends on a choice of basis of each of the two vector spaces. Now that we are studying operators, which map a vector space to itself, the emphasis is on using only one basis.

5.35 definition: *matrix of an operator, $\mathcal{M}(T)$*

Suppose $T \in \mathcal{L}(V)$. The *matrix of T* with respect to a basis $v_1, \dots, v_n$ of V is the n-by-n matrix

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix}$$

whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n.$$

The notation $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ is used if the basis is not clear from the context.

Operators have square matrices (meaning that the number of rows equals the number of columns), rather than the more general rectangular matrices that we considered earlier for linear maps.

If T is an operator on $\mathbf{F}^n$ and no basis is specified, assume that the basis in question is the standard one (where the k^{th} basis vector is 1 in the k^{th} slot and 0 in all other slots). You can then think of the k^{th} column of $\mathcal{M}(T)$ as T applied to the k^{th} basis vector, where we identify n-by-1 column vectors with elements of $\mathbf{F}^n$.

The k^{th} column of the matrix $\mathcal{M}(T)$ is formed from the coefficients used to write Tv_k as a linear combination of the basis $v_1, \dots, v_n$.

5.36 example: *matrix of an operator with respect to standard basis*

Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. Then the matrix of T with respect to the standard basis of $\mathbf{F}^3$ is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix},$$

as you should verify.

A central goal of linear algebra is to show that given an operator T on a finite-dimensional vector space V, there exists a basis of V with respect to which T has a reasonably simple matrix. To make this vague formulation a bit more precise, we might try to choose a basis of V such that $\mathcal{M}(T)$ has many 0's.

If V is a finite-dimensional complex vector space, then we already know enough to show that there is a basis of V with respect to which the matrix of T has 0's everywhere in the first column, except possibly the first entry. In other words, there is a basis of V with respect to which the matrix of T looks like

$$\begin{pmatrix} \lambda & & \\ 0 & * & \\ \vdots & & \\ 0 & & \end{pmatrix};$$

here $*$ denotes the entries in all columns other than the first column. To prove this, let λ be an eigenvalue of T (one exists by 5.19) and let v be a corresponding eigenvector. Extend v to a basis of V. Then the matrix of T with respect to this basis has the form above. Soon we will see that we can choose a basis of V with respect to which the matrix of T has even more 0's.

5.37 definition: *diagonal of a matrix*

The *diagonal* of a square matrix consists of the entries on the line from the upper left corner to the bottom right corner.

For example, the diagonal of the matrix

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$$

from Example 5.36 consists of the entries $2, 5, 8$, which are shown in red in the matrix above.

5.38 definition: *upper-triangular matrix*

A square matrix is called *upper triangular* if all entries below the diagonal are 0.

For example, the 3-by-3 matrix above is upper triangular.

Typically we represent an upper-triangular matrix in the form

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix};$$

the 0 in the matrix above indicates that all entries below the diagonal in this n-by-n matrix equal 0. Upper-triangular matrices can be considered reasonably simple—if n is large, then at least almost half the entries in an n-by-n upper-triangular matrix are 0.

We often use $*$ to denote matrix entries that we do not know or that are irrelevant to the questions being discussed.

The next result provides a useful connection between upper-triangular matrices and invariant subspaces.

5.39 *conditions for upper-triangular matrix*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then the following are equivalent.

(a) The matrix of T with respect to $v_1, \ldots, v_n$ is upper triangular.

(b) $\operatorname{span}(v_1, \ldots, v_k)$ is invariant under T for each $k = 1, \ldots, n$.

(c) $Tv_k \in \operatorname{span}(v_1, \ldots, v_k)$ for each $k = 1, \ldots, n$.

Proof First suppose (a) holds. To prove that (b) holds, suppose $k \in \{1, \ldots, n\}$. If $j \in \{1, \ldots, n\}$, then

$$Tv_j \in \operatorname{span}(v_1, \ldots, v_j)$$

because the matrix of T with respect to $v_1, \ldots, v_n$ is upper triangular. Because $\operatorname{span}(v_1, \ldots, v_j) \subseteq \operatorname{span}(v_1, \ldots, v_k)$ if $j \leq k$, we see that

$$Tv_j \in \operatorname{span}(v_1, \ldots, v_k)$$

for each $j \in \{1, \ldots, k\}$. Thus $\operatorname{span}(v_1, \ldots, v_k)$ is invariant under T, completing the proof that (a) implies (b).

Now suppose (b) holds, so $\operatorname{span}(v_1, \ldots, v_k)$ is invariant under T for each $k = 1, \ldots, n$. In particular, $Tv_k \in \operatorname{span}(v_1, \ldots, v_k)$ for each $k = 1, \ldots, n$. Thus (b) implies (c).

Now suppose (c) holds, so $Tv_k \in \operatorname{span}(v_1, \ldots, v_k)$ for each $k = 1, \ldots, n$. This means that when writing each Tv_k as a linear combination of the basis vectors $v_1, \ldots, v_n$, we need to use only the vectors $v_1, \ldots, v_k$. Hence all entries under the diagonal of $\mathcal{M}(T)$ are 0. Thus $\mathcal{M}(T)$ is an upper-triangular matrix, completing the proof that (c) implies (a).

We have shown that (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c) $\Longrightarrow$ (a), which shows that (a), (b), and (c) are equivalent. ■

The next result tells us that if $T \in \mathcal{L}(V)$ and with respect to some basis of V we have

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

then T satisfies a simple equation depending on $\lambda_1, \ldots, \lambda_n$.

5.40 *equation satisfied by operator with upper-triangular matrix*

Suppose $T \in \mathcal{L}(V)$ and V has a basis with respect to which T has an upper-triangular matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then

$$(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0.$$

Proof Let $v_1, \dots, v_n$ denote a basis of V with respect to which T has an upper-triangular matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Then $Tv_1 = \lambda_1 v_1$, which means that $(T - \lambda_1 I)v_1 = 0$, which implies that $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$ for $m = 1, \dots, n$ (using the commutativity of each $T - \lambda_j I$ with each $T - \lambda_k I$).

Note that $(T - \lambda_2 I)v_2 \in \operatorname{span}(v_1)$. Thus $(T - \lambda_1 I)(T - \lambda_2 I)v_2 = 0$ (by the previous paragraph), which implies that $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_2 = 0$ for $m = 2, \dots, n$ (using the commutativity of each $T - \lambda_j I$ with each $T - \lambda_k I$).

Note that $(T - \lambda_3 I)v_3 \in \operatorname{span}(v_1, v_2)$. Thus by the previous paragraph, $(T-\lambda_1 I)(T-\lambda_2 I)(T-\lambda_3 I)v_3 = 0$, which implies that $(T-\lambda_1 I) \cdots (T-\lambda_m I)v_3 = 0$ for $m = 3, \dots, n$ (using the commutativity of each $T - \lambda_j I$ with each $T - \lambda_k I$).

Continuing this pattern, we see that $(T - \lambda_1 I) \cdots (T - \lambda_n I)v_k = 0$ for each $k = 1, \dots, n$. Thus $(T - \lambda_1 I) \cdots (T - \lambda_n I)$ is the 0 operator because it is 0 on each vector in a basis of V. ■

Unfortunately no method exists for exactly computing the eigenvalues of an operator from its matrix. However, if we are fortunate enough to find a basis with respect to which the matrix of the operator is upper triangular, then the problem of computing the eigenvalues becomes trivial, as the next result shows.

5.41 *determination of eigenvalues from upper-triangular matrix*

Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper-triangular matrix.

Proof Suppose $v_1, \dots, v_n$ is a basis of V with respect to which T has an upper-triangular matrix

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Because $Tv_1 = \lambda_1 v_1$, we see that λ_1 is an eigenvalue of T.

Suppose $k \in \{2, \dots, n\}$. Then $(T - \lambda_k I)v_k \in \operatorname{span}(v_1, \dots, v_{k-1})$. Thus $T - \lambda_k I$ maps $\operatorname{span}(v_1, \dots, v_k)$ into $\operatorname{span}(v_1, \dots, v_{k-1})$. Because

$$\dim \operatorname{span}(v_1, \dots, v_k) = k \quad \text{and} \quad \dim \operatorname{span}(v_1, \dots, v_{k-1}) = k - 1,$$

this implies that $T - \lambda_k I$ restricted to $\operatorname{span}(v_1, \dots, v_k)$ is not injective (by 3.22). Thus there exists $v \in \operatorname{span}(v_1, \dots, v_k)$ such that $v \neq 0$ and $(T - \lambda_k I)v = 0$. Thus λ_k is an eigenvalue of T. Hence we have shown that every entry on the diagonal of $\mathcal{M}(T)$ is an eigenvalue of T.

To prove T has no other eigenvalues, let q be the polynomial defined by $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$. Then $q(T) = 0$ (by 5.40). Hence q is a polynomial multiple of the minimal polynomial of T (by 5.29). Thus every zero of the minimal polynomial of T is a zero of q. Because the zeros of the minimal polynomial of T are the eigenvalues of T (by 5.27), this implies that every eigenvalue of T is a zero of q. Hence the eigenvalues of T are all contained in the list $\lambda_1, \dots, \lambda_n$. ■

5.42 example: *eigenvalues via an upper-triangular matrix*

Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$. The matrix of T with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

Now 5.41 implies that the eigenvalues of T are 2, 5, and 8.

The next example illustrates 5.44: an operator has an upper-triangular matrix with respect to some basis if and only if the minimal polynomial of the operator is the product of polynomials of degree 1.

5.43 example: *whether T has an upper-triangular matrix can depend on* $\mathbf{F}$

Define $T \in \mathcal{L}(\mathbf{F}^4)$ by

$$T(z_1, z_2, z_3, z_4) = (-z_2, z_1, 2z_1 + 3z_3, z_3 + 3z_4).$$

Thus with respect to the standard basis of $\mathbf{F}^4$, the matrix of T is

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

You can ask a computer to verify that the minimal polynomial of T is the polynomial p defined by

$$p(z) = 9 - 6z + 10z^2 - 6z^3 + z^4.$$

First consider the case $\mathbf{F} = \mathbf{R}$. Then the polynomial p factors as

$$p(z) = (z^2 + 1)(z - 3)(z - 3),$$

with no further factorization of $z^2 + 1$ as the product of two polynomials of degree 1 with real coefficients. Thus 5.44 states that there does not exist a basis of $\mathbf{R}^4$ with respect to which T has an upper-triangular matrix.

Now consider the case $\mathbf{F} = \mathbf{C}$. Then the polynomial p factors as

$$p(z) = (z - i)(z + i)(z - 3)(z - 3),$$

where all factors above have the form $z - \lambda_k$. Thus 5.44 states that there is a basis of $\mathbf{C}^4$ with respect to which T has an upper-triangular matrix. Indeed, you can verify that with respect to the basis $(4-3i, -3-4i, -3+i, 1)$, $(4+3i, -3+4i, -3-i, 1)$, $(0, 0, 0, 1)$, $(0, 0, 1, 0)$ of $\mathbf{C}^4$, the operator T has the upper-triangular matrix

$$\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

5.44 *necessary and sufficient condition to have an upper-triangular matrix*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

Proof First suppose T has an upper-triangular matrix with respect to some basis of V. Let $\alpha_1, \ldots, \alpha_n$ denote the diagonal entries of that matrix. Define a polynomial $q \in \mathcal{P}(\mathbf{F})$ by

$$q(z) = (z - \alpha_1) \cdots (z - \alpha_n).$$

Then $q(T) = 0$, by 5.40. Hence q is a polynomial multiple of the minimal polynomial of T, by 5.29. Thus the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$ with $\{\lambda_1, \ldots, \lambda_m\} \subseteq \{\alpha_1, \ldots, \alpha_n\}$.

To prove the implication in the other direction, now suppose the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$. We will use induction on m. To get started, if $m = 1$ then $z - \lambda_1$ is the minimal polynomial of T, which implies that $T = \lambda_1 I$, which implies that the matrix of T (with respect to any basis of V) is upper triangular.

Now suppose $m > 1$ and the desired result holds for all smaller positive integers. Let

$$U = \operatorname{range}(T - \lambda_m I).$$

Then U is invariant under T [this is a special case of 5.18 with $p(z) = z - \lambda_m$]. Thus $T|_U$ is an operator on U.

If $u \in U$, then $u = (T - \lambda_m I)v$ for some $v \in V$ and

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0.$$

Hence $(z - \lambda_1) \cdots (z - \lambda_{m-1})$ is a polynomial multiple of the minimal polynomial of $T|_U$, by 5.29. Thus the minimal polynomial of $T|_U$ is the product of at most $m - 1$ terms of the form $z - \lambda_k$.

By our induction hypothesis, there is a basis $u_1, \ldots, u_M$ of U with respect to which $T|_U$ has an upper-triangular matrix. Thus for each $k \in \{1, \ldots, M\}$, we have (using 5.39)

5.45 $$Tu_k = (T|_U)(u_k) \in \operatorname{span}(u_1, \ldots, u_k).$$

Extend $u_1, \ldots, u_M$ to a basis $u_1, \ldots, u_M, v_1, \ldots, v_N$ of V. If $k \in \{1, \ldots, N\}$, then

$$Tv_k = (T - \lambda_m I)v_k + \lambda_m v_k.$$

The definition of U shows that $(T - \lambda_m I)v_k \in U = \operatorname{span}(u_1, \ldots, u_M)$. Thus the equation above shows that

5.46 $$Tv_k \in \operatorname{span}(u_1, \ldots, u_M, v_1, \ldots, v_k).$$

From 5.45 and 5.46, we conclude (using 5.39) that T has an upper-triangular matrix with respect to the basis $u_1, \ldots, u_M, v_1, \ldots, v_N$ of V, as desired. ■

The set of numbers $\{\lambda_1, \ldots, \lambda_m\}$ from the previous result equals the set of eigenvalues of T (because the set of zeros of the minimal polynomial of T equals the set of eigenvalues of T, by 5.27), although the list $\lambda_1, \ldots, \lambda_m$ in the previous result may contain repetitions.

In Chapter 8 we will improve even the wonderful result below; see 8.37 and 8.46.

5.47 *if* $\mathbf{F} = \mathbf{C}$, *then every operator on* V *has an upper-triangular matrix*

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V.

Proof The desired result follows immediately from 5.44 and the second version of the fundamental theorem of algebra (see 4.13). ■

For an extension of the result above to two operators S and T such that

$$ST = TS,$$

see 5.80. Also, for an extension to more than two operators, see Exercise 9(b) in Section 5E.

Caution: If an operator $T \in \mathcal{L}(V)$ has a upper-triangular matrix with respect to some basis $v_1, \ldots, v_n$ of V, then the eigenvalues of T are exactly the entries on the diagonal of $\mathcal{M}(T)$, as shown by 5.41, and furthermore v_1 is an eigenvector of T. However, $v_2, \ldots, v_n$ need not be eigenvectors of T. Indeed, a basis vector v_k is an eigenvector of T if and only if all entries in the k^{th} column of the matrix of T are 0, except possibly the k^{th} entry.

You may recall from a previous course that every matrix of numbers can be changed to a matrix in what is called row echelon form. If one begins with a square matrix, the matrix in row echelon form will be an upper-triangular matrix. Do not confuse this upper-triangular matrix with the upper-triangular matrix of an operator with respect to some basis whose existence is proclaimed by 5.47 (if $\mathbf{F} = \mathbf{C}$)—there is no connection between these upper-triangular matrices.

The row echelon form of the matrix of an operator does not give us a list of the eigenvalues of the operator. In contrast, an upper-triangular matrix with respect to some basis gives us a list of all the eigenvalues of the operator. However, there is no method for computing exactly such an upper-triangular matrix, even though 5.47 guarantees its existence if $\mathbf{F} = \mathbf{C}$.

Exercises 5C

1 Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and T^2 has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some basis of V.

2 Suppose A and B are upper-triangular matrices of the same size, with $\alpha_1, \ldots, \alpha_n$ on the diagonal of A and $\beta_1, \ldots, \beta_n$ on the diagonal of B.

(a) Show that $A + B$ is an upper-triangular matrix with $\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n$ on the diagonal.

(b) Show that AB is an upper-triangular matrix with $\alpha_1\beta_1, \ldots, \alpha_n\beta_n$ on the diagonal.

The results in this exercise are used in the proof of 5.81.

3 Suppose $T \in \mathcal{L}(V)$ is invertible and $v_1, \ldots, v_n$ is a basis of V with respect to which the matrix of T is upper triangular, with $\lambda_1, \ldots, \lambda_n$ on the diagonal. Show that the matrix of T^{-1} is also upper triangular with respect to the basis $v_1, \ldots, v_n$, with

$$\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}$$

on the diagonal.

4 Give an example of an operator whose matrix with respect to some basis contains only 0's on the diagonal, but the operator is invertible.

This exercise and the exercise below show that 5.41 fails without the hypothesis that an upper-triangular matrix is under consideration.

5 Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.

6 Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that if $k \in \{1, \ldots, \dim V\}$, then V has a k-dimensional subspace invariant under T.

7 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$.

(a) Prove that there exists a unique monic polynomial p_v of smallest degree such that $p_v(T)v = 0$.

(b) Prove that the minimal polynomial of T is a polynomial multiple of p_v.

8 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and there exists a nonzero vector $v \in V$ such that $T^2v + 2Tv = -2v$.

(a) Prove that if $\mathbf{F} = \mathbf{R}$, then there does not exist a basis of V with respect to which T has an upper-triangular matrix.

(b) Prove that if $\mathbf{F} = \mathbf{C}$ and A is an upper-triangular matrix that equals the matrix of T with respect to some basis of V, then $-1 + i$ or $-1 - i$ appears on the diagonal of A.

9 Suppose B is a square matrix with complex entries. Prove that there exists an invertible square matrix A with complex entries such that $A^{-1}BA$ is an upper-triangular matrix.

10 Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Show that the following are equivalent.

(a) The matrix of T with respect to $v_1, \ldots, v_n$ is lower triangular.

(b) $\operatorname{span}(v_k, \ldots, v_n)$ is invariant under T for each $k = 1, \ldots, n$.

(c) $Tv_k \in \operatorname{span}(v_k, \ldots, v_n)$ for each $k = 1, \ldots, n$.

A square matrix is called ***lower triangular*** *if all entries above the diagonal are* 0*.*

11 Suppose $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional. Prove that if $T \in \mathcal{L}(V)$, then there exists a basis of V with respect to which T has a lower-triangular matrix.

12 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V, and U is a subspace of V that is invariant under T.

(a) Prove that $T|_U$ has an upper-triangular matrix with respect to some basis of U.

(b) Prove that the quotient operator T/U has an upper-triangular matrix with respect to some basis of V/U.

The quotient operator T/U *was defined in Exercise 38 in Section 5A.*

13 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose there exists a subspace U of V that is invariant under T such that $T|_U$ has an upper-triangular matrix with respect to some basis of U and also T/U has an upper-triangular matrix with respect to some basis of V/U. Prove that T has an upper-triangular matrix with respect to some basis of V.

14 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an upper-triangular matrix with respect to some basis of V if and only if the dual operator T' has an upper-triangular matrix with respect to some basis of the dual space V'.