7.5: Upper Triangular Matrices

As before, let \( V \) be a complex vector space.

Let \( \mathcal{L}(V, V) \) and \( (v_1, \ldots, v_n) \) be a basis for \( V \). Recall that we can associate a matrix \( M(T) \in \mathbb{C}^{n \times n} \) to the operator \( T \). By Theorem 7.4.1, we know that \( T \) has at least one eigenvalue, say \( \lambda \in \mathbb{C} \). Let \( v_1 \neq 0 \) be an eigenvector corresponding to \( \lambda \). By the Basis Extension Theorem, we can extend the list \( (v_1) \) to a basis of \( V \). Since \( Tv_1 = \lambda v_1 \), the first column of \( M(T) \) with respect to this basis is

\[
\begin{bmatrix}
\lambda \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

What we will show next is that we can find a basis of \( V \) such that the matrix \( M(T) \) is upper triangular.

Definition 7.5.1: Upper Trianglar Matrix

A matrix \( A = (a_{ij}) \in \mathbb{F}^{n \times n} \) is called upper triangular if \( a_{ij} = 0 \) for \( i > j \).

Schematically, an upper triangular matrix has the form

\[
\begin{bmatrix}
* & * & \cdots & \cdots & * \\
& * & \cdots & \cdots & * \\
& & \ddots & \cdots & * \\
& & & \ddots & * \\
& & & & *
\end{bmatrix}
\]

where the entries \( (\ast) \) can be anything and every entry below the main diagonal is zero.

Here are two reasons why having an operator \( T \) represented by an upper triangular matrix can be quite convenient:

1. the eigenvalues are on the diagonal (as we will see later);
2. it is easy to solve the corresponding system of linear equations by back substitution (as discussed in Section A.3).

The next proposition tells us what upper triangularity means in terms of linear operators and invariant subspaces.

Proposition 7.5.2

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

1. the matrix \(M(T)\) with respect to the basis \((v_1, \ldots, v_n)\) is upper triangular;
2. \(Tv_k \in \text{Span}(v_1, \ldots, v_k)\) for each \(k=1,2,\ldots,n\);
3. \(\text{Span}(v_1, \ldots, v_k)\) is invariant under \(T\) for each \(k=1,2,\ldots,n\).

Proof

The equivalence of Condition~1 and Condition~2 follows easily from the definition since Condition~2 implies that the matrix elements below the diagonal are zero.

Obviously, Condition~3 implies Condition~2. To show that Condition~2 implies Condition~3, note that any vector \(v \in \text{Span}(v_1, \ldots, v_k)\) can be written as \(v = a_1v_1 + \cdots + a_kv_k\). Applying \((T)\), we obtain

\[
Tv = a_1Tv_1 + \cdots + a_kTv_k \in \text{Span}(v_1, \ldots, v_k)
\]

since, by Condition~2, each \(Tv_j \in \text{Span}(v_1, \ldots, v_j)\subset \text{Span}(v_1, \ldots, v_k)\) for \(j=1,2,\ldots,k\) and since the span is a subspace of \(V\).

\(\square\)

The next theorem shows that complex vector spaces indeed have some basis for which the matrix of a given operator is upper triangular.

Theorem 7.5.3

Let \(V\) be a finite-dimensional vector space over \(\mathbb{C}\) and \(T \in \mathcal{L}(V,V)\). Then there exists a basis \(B\) for \(V\) such that \(M(T)\) is upper triangular with respect to \(B\).

Proof

We proceed by induction on \(\dim(V)\). If \(\dim(V)=1\), then there is nothing to prove.

Hence, assume that \(\dim(V)>1\) and that we have proven the result for all \(T \in \mathcal{L}(W,W)\), where \(W\) is a complex vector space with \(\dim(W)\le n-1\). By Theorem 7.4.1, \(T\) has at least one eigenvalue \(\lambda\).

Define

\[
U = \text{range}(T-\lambda I)
\]
and note that

1. \(\dim(U) < \dim(V) = n\) since \(\lambda\) is an eigenvalue of \(T\) and hence \(T - \lambda I\) is not surjective;
2. \(U\) is an invariant subspace of \(T\) since, for all \(u \in U\), we have

\[
Tu = (T - \lambda I) u + \lambda u,
\]

which implies that \(Tu \in U\) since \((T - \lambda I) u \in \text{range}(T - \lambda I) = U\) and \(\lambda u \in U\).

Therefore, we may consider the operator \((S = T|_U)\), which is the operator obtained by restricting \(T\) to the subspace \(U\). By the induction hypothesis, there exists a basis \(\{(u_1, \ldots, u_m)\}\) of \(U\) with \(m \leq n-1\) such that \((M(S))\) is upper triangular with respect to \(\{(u_1, \ldots, u_m, u_{m+1}, \ldots, u_n)\}\). This means that

\[
Tu_j = Su_j \in \text{span}(u_1, \ldots, u_j), \quad \text{for all } j = 1, 2, \ldots, m.
\]

Extend this to a basis \(\{(u_1, \ldots, u_m, v_1, \ldots, v_k)\}\) of \(V\). Then

\[
Tv_j = (T - \lambda I)v_j + \lambda v_j, \quad \text{for all } j = 1, 2, \ldots, k.
\]

Since \((T - \lambda I)v_j \in \text{range}(T - \lambda I) = U = \text{span}(u_1, \ldots, u_m)\), we have that

\[
Tv_j \in \text{span}(u_1, \ldots, u_m, v_1, \ldots, v_j), \quad \text{for all } j = 1, 2, \ldots, k.
\]

Hence, \(T\) is upper triangular with respect to the basis \(\{(u_1, \ldots, u_m, v_1, \ldots, v_k)\}\).

The following are two very important facts about upper triangular matrices and their associated operators.

Proposition 7.5.4

Suppose \(T \in \mathcal{L}(V, V)\) is a linear operator and that \((M(T))\) is upper triangular with respect to some basis of \(V\).

1. \(T\) is invertible if and only if all entries on the diagonal of \((M(T))\) are nonzero.
2. The eigenvalues of \(T\) are precisely the diagonal elements of \((M(T))\).

Proof of Proposition 7.5.4, Part 1

Let \(\{(v_1, \ldots, v_n)\}\) be a basis of \(V\) such that

\[
\begin{align*}
M(T) &= \begin{bmatrix}
\lambda_1 & & \\
& \ddots & \\
0 & & \lambda_n
\end{bmatrix}, \\
&= \lambda_1 & & \\
& & \ddots & \\
&0 & & \lambda_n
\end{align*}
\]
is upper triangular. The claim is that \((T)\) is invertible if and only if \(\lambda_k \neq 0\) for all \(k=1,2,\ldots,n\). Equivalently, this can be reformulated as follows: \((T)\) is not invertible if and only if \(\lambda_k = 0\) for at least one \(k\in\{1,2,\ldots,n\}\).

Suppose \(\lambda_k = 0\). We will show that this implies the non-invertibility of \((T)\). If \(k=1\), this is obvious since then \((Tv_1 = 0)\), which implies that \((v_1 \in \ker(T))\) so that \((T)\) is not injective and hence not invertible. So assume that \((k>1)\). Then

\[
Tv_j \in \Span(v_1,\ldots,v_{k-1}), \quad \text{for all } j \leq k,
\]

since \((T)\) is upper triangular and \(\lambda_k = 0\). Hence, we may define \((S = T|_{\Span(v_1,\ldots,v_k)})\) to be the restriction of \((T)\) to the subspace \(\Span(v_1,\ldots,v_k)\) so that

\[
[S: \Span(v_1,\ldots,v_k) \to \Span(v_1,\ldots,v_{k-1}).]
\]

The linear map \((S)\) is not injective since the dimension of the domain is larger than the dimension of its codomain, i.e.,

\[
\dim(\Span(v_1,\ldots,v_k)) = k > k-1 = \dim(\Span(v_1,\ldots,v_{k-1})).
\]

Hence, there exists a vector \((v \in \Span(v_1,\ldots,v_k))\) such that \((Sv = Tv = 0)\). This implies that \((T)\) is also not injective and therefore also not invertible.

Now suppose that \((T)\) is not invertible. We need to show that at least one \(\lambda_k = 0\). The linear map \((T)\) not being invertible implies that \((T)\) is not injective. Hence, there exists a vector \((v \in V)\) such that \((Tv = 0)\), and we can write

\[
\begin{align*}
v &= a_1 v_1 + \cdots + a_k v_k \\
\end{align*}
\]

for some \((k)\), where \((a_k \neq 0)\). Then

\[
\begin{align*}
0 &= Tv = (a_1 Tv_1 + \cdots + a_{k-1}Tv_{k-1}) + a_k Tv_k. \label{eq:expansion}
\end{align*}
\]

Since \((T)\) is upper triangular with respect to the basis \(((v_1,\ldots,v_n))\), we know that \((a_1 Tv_1 + \cdots + a_{k-1}Tv_{k-1}) \in \Span(v_1,\ldots,v_{k-1})\)). Hence, Equation \ref{eq:expansion} shows that \((Tv_k \in \Span(v_1,\ldots,v_{k-1}))\), which implies that \((\lambda_k = 0)\).

\(\square\)

Proof of Proposition 7.5.4, Part 2.

Recall that \((\lambda \in \mathbb{F})\) is an eigenvalue of \((T)\) if and only if the operator \((T-\lambda I)\) is not invertible. Let \((\{v_1,\ldots,v_n\})\) be a basis such that \((M(T))\) is upper triangular. Then
\begin{equation*}
M(T-\lambda I) = \begin{bmatrix} \lambda_1-\lambda & \& \ast \\
& \ddots & \\
0 & \& \lambda_n-\lambda 
\end{bmatrix}.
\end{equation*}

Hence, by Proposition 7.5.4(1), \((T-\lambda I)\) is not invertible if and only if \((\lambda=\lambda_k)\) for some \(k\).