11.4 Applications of the Spectral Theorem: diagonalization

Let \((e=(e_1,\ldots,e_n))\) be a basis for an \((n)\)-dimensional vector space \((V)\), and let \((T\in \mathcal{L}(V))\). In this section we denote the matrix \((M(T))\) of \((T)\) with respect to basis \((e)\) by \([T]_e\). This is done to emphasize the dependency on the basis \((e)\).

In other words, we have that

\[
[Tv]_e = [T]_e [v]_e, \quad \text{for all } v \in V,
\]

where

\[
[v]_e = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}
\]

is the coordinate vector for \((v= v_1 e_1 + \cdots + v_n e_n)\) with \((v_i \in \mathbb{F})\).

The operator \((T)\) is diagonalizable if there exists a basis \((e)\) such that \([T]_e\) is diagonal, i.e., if there exist \((\lambda_1,\ldots,\lambda_n \in \mathbb{F})\) such that

\[
[T]_e = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}
\]

The scalars \((\lambda_1,\ldots,\lambda_n)\) are necessarily eigenvalues of \((T)\), and \((e_1,\ldots,e_n)\) are the corresponding eigenvectors. We summarize this in the following proposition.
Proposition 11.4.1. \( (T\in \mathcal{L}(V)) \) is diagonalizable if and only if there exists a basis \( ((e_1, \ldots, e_n)) \) consisting entirely of eigenvectors of \( (T) \).

We can reformulate this proposition using the change of basis transformations as follows. Suppose that \( (e) \) and \( (f) \) are bases of \( (V) \) such that \( ([T]_e) \) is diagonal, and let \( (S) \) be the change of basis transformation such that \( ([v]_e=S[v]_f) \). Then \( ([S][T][S]^{-1}]=[T]_e \) is diagonal.

Proposition 11.4.2. \( (T\in \mathcal{L}(V)) \) is diagonalizable if and only if there exists an invertible matrix \( (S\in \mathbb{F}^{n\times n}) \) such that

\[
\begin{bmatrix} \lambda_1 & \ldots & 0 \\
0 & \ddots & \vdots \\
0 & \ldots & \lambda_n 
\end{bmatrix},
\]

where \( ([T]_f) \) is the matrix for \( (T) \) with respect to a given arbitrary basis \( (f=(f_1, \ldots, f_n)) \).

On the other hand, the Spectral Theorem tells us that \( (T) \) is diagonalizable with respect to an orthonormal basis if and only if \( (T) \) is normal. Recall that

\[
\begin{bmatrix} A^* \end{bmatrix}_f = \begin{bmatrix} A \end{bmatrix}_f^*,
\]

for any orthonormal basis \( (f) \) of \( (V) \). As before,

\[
A^* = (\overline{a}_{ji})_{ij=1}^n, \quad \text{for } A=(a_{ij})_{i,j=1}^n,
\]

is the conjugate transpose of the matrix \( (A) \). When \( (\mathbb{F}=\mathbb{R}) \), note that \( (A^*=A^\dagger) \) is just the transpose of the matrix, where \( (A^\dagger=(a_{ji})_{i,j=1}^n) \).

The change of basis transformation between two orthonormal bases is called unitary in the complex case and orthogonal in the real case. Let \( (e=(e_1, \ldots, e_n)) \) and \( (f=(f_1, \ldots, f_n)) \) be two orthonormal bases of \( (V) \), and let \( (U) \) be the change of basis matrix such that \( ([v]_f=U[v]_e) \), for all \( (v\in V) \). Then

\[
\begin{bmatrix} \text{inner}\{e_i, e_j\} = \delta_{ij} = \text{inner}\{f_i, f_j\} = \text{inner}\{Ue_i, Ue_j\} \end{bmatrix}_e
\]

Since this holds for the basis \( (e) \), it follows that \( (U) \) is unitary if and only if

\[
\begin{bmatrix} \text{inner}\{Uv, Uw\} = \text{inner}\{v, w\} \quad \text{for all } (v, w\in V) \end{bmatrix}_e \tag{11.4.1}
\]

This means that unitary matrices preserve the inner product. Operators that preserve the inner product are often also called isometries. Orthogonal matrices also define isometries.

By the definition of the adjoint, \( (\text{inner}\{Uv, Uw\} = \text{inner}\{v, U^*Uw\}) \), and so Equation 11.4.1 implies that isometries are characterized by the property
\begin{equation*} U^*U &= I, \quad \text{for the unitary case}, \\ O^T O &= I, \quad \text{for the orthogonal case.} \end{equation*}

The equation \((U^*U=I)\) implies that \((U^{-1}=U^*)\). For finite-dimensional inner product spaces, the left inverse of an operator is also the right inverse, and so

\begin{array}{c} UU^* = I \quad \text{if and only if} \quad U^*U=I, \quad \text{if and only if} \quad O^T O = I \quad \text{if and only if} \quad O^* O = I, \tag{11.4.2} \end{array}

It is easy to see that the columns of a unitary matrix are the coefficients of the elements of an orthonormal basis with respect to another orthonormal basis. Therefore, the columns are orthonormal vectors in \(\mathbb{C}^n\) (or in \(\mathbb{R}^n\) in the real case). By Condition (11.4.2), this is also true for the rows of the matrix.

The Spectral Theorem tells us that \(T \in \mathcal{L}(V)\) is normal if and only if \([T]_e\) is diagonal with respect to an orthonormal basis \((e_i)\) for \((V)\), i.e., if there exists a unitary matrix \((U)\) such that

\begin{equation*} UTU^* = \begin{bmatrix} \lambda_1 \& \ddots \& 0 \\ & \ddots & \vdots \\ 0 \& \cdots \& \lambda_n \end{bmatrix}. \end{equation*}

Conversely, if a unitary matrix \((U)\) exists such that \((UTU^*=D)\) is diagonal, then

\begin{equation*} TT^* - T^*T = U^*(D\overline{D}-\overline{D}D)U=0 \end{equation*}

since diagonal matrices commute, and hence \((T)\) is normal. Let us summarize some of the definitions that we have seen in this section.

**Definition 11.4.3.** Given a square matrix \(A \in \mathbb{F}^{n \times n}\), we call

1. **symmetric** if \((A = A^T)\).
2. **Hermitian** if \((A = A^*)\).
3. **orthogonal** if \((AA^T = I)\).
4. **unitary** if \((AA^* = I)\).

Note that every type of matrix in Definition 11.4.3 is an example of a normal operator. An example of a normal operator \((N)\) that is neither Hermitian nor unitary is
\[ N = i \left[ \begin{array}{cc} -1 & -1 \\ -1 & 1 \end{array} \right] \]

You can easily verify that \( NN^* = N^*N \) and that \( iN \) is symmetric (not Hermitian).

**Example 11.4.4.** Consider the matrix
\[
A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}
\]
from Example 11.1.5~\ref{ex:hermitian}. To unitarily diagonalize \( A \), we need to find a unitary matrix \( U \) and a diagonal matrix \( D \) such that \( A=UDU^{-1} \). To do this, we need to first find a basis for \( \mathbb{C}^2 \) that consists entirely of orthonormal eigenvectors for the linear map \( T \) in \( \mathcal{L}(\mathbb{C}^2) \) defined by \( Tv=Av \), for all \( v \) in \( \mathbb{C}^2 \).

To find such an orthonormal basis, we start by finding the eigenspaces of \( T \). We already determined that the eigenvalues of \( T \) are \( \lambda_1=1 \) and \( \lambda_2=4 \), so \( D = \begin{bmatrix} 1&0 \\ 0&4 \end{bmatrix} \). It follows that
\[
A = UDU^{-1} = \begin{bmatrix} -1-i/\sqrt{3} & 1+i/\sqrt{6} \\ 1/\sqrt{3} & 2/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1&0 \\ 0&4 \end{bmatrix} \begin{bmatrix} -1+i/\sqrt{3} & 1/\sqrt{3} \\ 1-i/\sqrt{6} & 2/\sqrt{6} \end{bmatrix}.
\]
As an application, note that such diagonal decomposition allows us to easily compute powers and the exponential of matrices. Namely, if \( A = U D U^{-1} \) with \( D \) diagonal, then we have

\[
\begin{align*}
A^n &= (UDU^{-1})^n = U D^n U^{-1}, \\
\exp(A) &= \sum_{k=0}^{\infty} \frac{1}{k!} A^k \\
&= U \left( \sum_{k=0}^{\infty} \frac{1}{k!} D^k \right) U^{-1} = U \exp(D) U^{-1}.
\end{align*}
\]

**Example 11.4.5.** Continuing Example 11.4.4,

\[
\begin{align*}
A^2 &= (UDU^{-1})^2 = UD^2 U^{-1} = U \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix} U^* = \begin{bmatrix} 6 & 5+5i \\ 5-5i & 11 \end{bmatrix}, \\
A^n &= (UDU^{-1})^n = UD^n U^{-1} = U \begin{bmatrix} 1 & 0 \\ 0 & 2^{2n} \end{bmatrix} U^* = \begin{bmatrix} \frac{2}{3}(1+2^{n-1}) & \frac{1+i}{3}(-1+2^{2n}) \\ \frac{1-i}{3}(-1+2^{2n}) & \frac{1}{3}(1+2^{2n+1}) \end{bmatrix}, \\
\exp(A) &= U \exp(D) U^{-1} = U \begin{bmatrix} e & 0 \\ 0 & e^4 \end{bmatrix} U^{-1} = \frac{1}{3} \begin{bmatrix} 2e+e^4 & e^4-e+i(e^4-e) \\ e^4-e+i(e-e^4) & e+2e^4 \end{bmatrix}.
\end{align*}
\]

**Contributors**

- Isaiah Lankham, Mathematics Department at UC Davis
- Bruno Nachtergaele, Mathematics Department at UC Davis
- Anne Schilling, Mathematics Department at UC Davis

Both hardbound and softbound versions of this textbook are available online at WorldScientific.com.