before, we have $v = \sum_{j=1}^n \langle v, f_j\rangle f_j$. Comparing this with $v = \sum_{i=1}^n \langle v, e_i\rangle e_i$, we find that

$$v = \cancel{\sum_{i,j=1}^n \langle \langle v, e_j\rangle e_j, f_i\rangle f_i = \sum_{i=1}^n \left(\sum_{j=1}^n \langle e_j, f_i\rangle \langle v, e_j\rangle \right) f_i} \quad \sum_i \langle v, e_i \rangle e_i$$

$$= \sum_i \sum_j \langle \langle v, f_j \rangle f_j, e_i \rangle e_i = \sum_i \left(\sum_j \langle f_j, e_i \rangle \langle v, f_j \rangle \right) e_i$$

Hence,

$$[v]_e = S[v]_f$$

where

$$S = (s_{ij})_{i,j=1}^n \qquad \text{with } s_{ij} = \cancel{\langle e_j, f_i\rangle} \; \langle f_j, e_i \rangle$$

The j^{th} column of S is given by the coefficients of the expansion of f_j in terms of the basis e. ~~$f = (f_1, \ldots, f_n)$.~~ The matrix S describes a linear map in $\mathcal{L}(\mathbb{F}^n)$; ~~which~~ S is called the **change of basis** ~~**transformation**~~ matrix. *

We may also interchange the role of bases e and f. In this case, we obtain the matrix $R = (r_{ij})_{i,j=1}^n$, where

$$r_{ij} = \langle e_j, f_i\rangle.$$

Then, by the uniqueness of the expansion in a basis, we obtain

$$[v]_f = R[v]_e$$

so that

$$RS[v]_f = [v]_f, \qquad \text{for all } v \in V.$$

Since this equation is true for all $[v]_f \in \mathbb{F}^n$, it follows that either $RS = I$ or $R = S^{-1}$. In particular, S and R are invertible. We can also check this explicitly by using the properties of orthonormal bases. Namely,

$$\begin{aligned}
(SR)_{ij} &= \sum_{k=1}^n s_{ik} r_{kj} = \sum_{k=1}^n \langle f_k, e_i\rangle \langle e_j, f_k\rangle \\
&= \sum_{k=1}^n \langle e_j, f_k\rangle \overline{\langle e_i, f_k\rangle} = \langle [e_j]_f, [e_i]_f\rangle_{\mathbb{F}^n} = \delta_{ij}.
\end{aligned}$$

Matrix S (and similarly also R) has the interesting property that its columns are orthonormal to one another. This follows from the fact that the columns are the coordinates of

* Change of basis from e to f.

orthonormal vectors with respect to another orthonormal basis. A similar statement holds for the rows of S (and similarly also R).

Example 10.2.2. Let $V = \mathbb{C}^2$, and choose the orthonormal bases $e = (e_1, e_2)$ and $f = (f_1, f_2)$ with

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$f_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad f_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then

$$R\,\cancel{S} = \begin{bmatrix} \langle e_1, f_1 \rangle & \langle e_2, f_1 \rangle \\ \langle e_1, f_2 \rangle & \langle e_2, f_2 \rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

and

$$S\,R = \begin{bmatrix} \langle f_1, e_1 \rangle & \langle f_2, e_1 \rangle \\ \langle f_1, e_2 \rangle & \langle f_2, e_2 \rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

One can then check explicitly that indeed

$$SR\ \cancel{RS} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

So far we have only discussed how the coordinate vector of a given vector $v \in V$ changes under the change of basis from e to f. The next question we can ask is how the matrix $M(T)$ of an operator $T \in \mathcal{L}(V)$ changes if we change the basis. Let A be the matrix of T with respect to the basis $e = (e_1, \ldots, e_n)$, and let B be the matrix for T with respect to the basis $f = (f_1, \ldots, f_n)$. How do we determine B from A? Note that

$$[Tv]_e = A[v]_e$$

so that

$$S\,[Tv]_f = \cancel{S}[Tv]_e = \cancel{S}A[v]_e = \cancel{S}A\overset{S}{\cancel{R}}[v]_f \cancel{= SAS^{-1}[v]_f}. \Rightarrow [Tv]_f = S^{-1}AS[v]_f$$

This implies that

$$B = \cancel{SAS^{-1}}.\ S^{-1}AS\,.$$

Example 10.2.3. Continuing Example 10.2.2, let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

be the matrix of a linear operator with respect to the basis e. Then the matrix B with respect to the basis f is given by

$$B = \overset{S^{-1}AS}{\cancel{SAS^{-1}}} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

Exercises for Chapter 10

Calculational Exercises

1. Consider $\mathbb{R}^3$ with two orthonormal bases: the canonical basis $e = (e_1, e_2, e_3)$ and the basis $f = (f_1, f_2, f_3)$, where

$$f_1 = \frac{1}{\sqrt{3}}(1,1,1),\ f_2 = \frac{1}{\sqrt{6}}(1,-2,1),\ f_3 = \frac{1}{\sqrt{2}}(1,0,-1).$$

Find the matrix, S, of the change of basis transformation such that

$$[v]_{\cancel{f}\atop e} = S[v]_{\cancel{e}\atop f}, \quad \text{for all } v \in \mathbb{R}^3,$$

where $[v]_b$ denotes the column vector of v with respect to the basis b.

2. Let $v \in \mathbb{C}^4$ be the vector given by $v = (1, i, -1, -i)$. Find the matrix (with respect to the canonical basis on $\mathbb{C}^4$) of the orthogonal projection $P \in \mathcal{L}(\mathbb{C}^4)$ such that

$$\operatorname{null}(P) = \{v\}^{\perp}.$$

3. Let U be the subspace of $\mathbb{R}^3$ that coincides with the plane through the origin that is perpendicular to the vector $n = (1,1,1) \in \mathbb{R}^3$.

   (a) Find an orthonormal basis for U.

   (b) Find the matrix (with respect to the canonical basis on $\mathbb{R}^3$) of the orthogonal projection $P \in \mathcal{L}(\mathbb{R}^3)$ onto U, i.e., such that $\operatorname{range}(P) = U$.

4. Let $V = \mathbb{C}^4$ with its standard inner product. For $\theta \in \mathbb{R}$, let

$$v_\theta = \begin{pmatrix} 1 \\ e^{i\theta} \\ e^{2i\theta} \\ e^{3i\theta} \end{pmatrix} \in \mathbb{C}^4.$$

Find the canonical matrix of the orthogonal projection onto the subspace $\{v_\theta\}^{\perp}$.

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

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

where

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

is the coordinate vector for $v = v_1e_1 + \cdots + v_ne_n$ 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 $[v]_{\cancel{e}} = S[v]_{\cancel{f}}$. Then $\cancel{S[T]_f S^{-1}} = [T]_e$ is diagonal. *
f e $S^{-1}[T]_f S$

Proposition 11.4.2. *$T \in \mathcal{L}(V)$ is diagonalizable if and only if there exists an invertible*

* Here change of basis from f to e.

matrix $S \in \mathbb{F}^{n\times n}$ such that

$$S^{-1}[T]_f S \; \cancel{S[T]_f S^{-1}} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix},$$

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

$$[T^*]_f = [T]_f^*$$

for any orthonormal basis f of V. As before,

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

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

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

$$\langle e_i, e_j \rangle = \delta_{ij} = \langle f_i, f_j \rangle = \langle Ue_i, Ue_j \rangle.$$

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

$$\langle Uv, Uw \rangle = \langle v, w \rangle \quad \text{for all } v, w \in V. \tag{11.1}$$

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, $\langle Uv, Uw \rangle = \langle v, U^*Uw \rangle$, and so Equation 11.1 implies that isometries are characterized by the property

$$\begin{aligned} U^*U &= I, && \text{for the unitary case,} \\ O^TO &= I, && \text{for the orthogonal case.} \end{aligned}$$

* Here U is the operator associated with matrix U.

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

$$\begin{aligned} UU^* = I &\quad \text{if and only if} \quad U^*U = I, \\ OO^T = I &\quad \text{if and only if} \quad O^TO = I. \end{aligned} \tag{11.2}$$

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 (11.2), 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 for V, i.e., if there exists a unitary matrix U such that

$$U^*[T]_f U \cancel{UTU^*} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

Conversely, if a unitary matrix U exists such that $\cancel{UTU^*}$ $U^*[T]_f U$ $= D$ is diagonal, then

$$[TT^* - T^*T]_f = U^{\cancel{*}}(D\overline{D} - \overline{D}D)U^{*} = 0$$

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.

4. Let $r \in \mathbb{R}$ and let $T \in \mathcal{L}(\mathbb{C}^2)$ be the linear map with canonical matrix

$$T = \begin{pmatrix} 1 & -1 \\ -1 & r \end{pmatrix}.$$

   (a) Find the eigenvalues of T.

   (b) Find an orthonormal basis of $\mathbb{C}^2$ consisting of eigenvectors of T.

   (c) Find a unitary matrix U such that $\cancel{UTU^*}$ U^*TU is diagonal.

5. Let A be the complex matrix given by:

$$A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -1 & -1+i \\ 0 & -1-i & 0 \end{bmatrix}$$

   (a) Find the eigenvalues of A.

   (b) Find an orthonormal basis of eigenvectors of A.

   (c) Calculate $|A| = \sqrt{A^*A}$.

   (d) Calculate e^A.

6. Let $\theta \in \mathbb{R}$, and let $T \in \mathcal{L}(\mathbb{C}^2)$ have canonical matrix

$$M(T) = \begin{pmatrix} 1 & e^{i\theta} \\ e^{-i\theta} & -1 \end{pmatrix}.$$

   (a) Find the eigenvalues of T.

   (b) Find an orthonormal basis for $\mathbb{C}^2$ that consists of eigenvectors for T.

Proof-Writing Exercises

1. Prove or give a counterexample: The product of any two self-adjoint operators on a finite-dimensional vector space is self-adjoint.

2. Prove or give a counterexample: Every unitary matrix is invertible.