24. Eigenvectors, spectral theorems

1. *Eigenvectors, eigenvalues*

Let k be a field, not necessarily algebraically closed.

Let T be a k-linear endomorphism of a k-vectorspace V to itself, meaning, as usual, that

$$T(v+w) = Tv + TW \qquad \text{and} \qquad T(cv) = c \cdot Tv$$

for $v, w \in V$ and $c \in k$. The collection of all such T is denoted $\mathrm{End}_k(V)$, and is a vector space over k with the natural operations

$$(S+T)(v) = Sv + Tv \qquad (cT)(v) = c \cdot Tv$$

A vector $v \in V$ is an **eigenvector** for T with **eigenvalue** $c \in k$ if

$$T(v) = c \cdot v$$

or, equivalently, if

$$(T - c \cdot \mathrm{id}_V)\, v = 0$$

A vector v is a **generalized eigenvector** of T with **eigenvalue** $c \in k$ if, for some integer $\ell \geq 1$

$$(T - c \cdot \mathrm{id}_V)^\ell\, v = 0$$

We will often suppress the id_V notation for the identity map on V, and just write c for the scalar operator $c \cdot \text{id}_V$. The collection of all λ-eigenvectors for T is the **λ-eigenspace** for T on V, and the collection of all generalized λ-eigenvectors for T is the **generalized λ-eigenspace** for T on V.

[1.0.1] Proposition: Let $T \in \text{End}_k(V)$. For fixed $\lambda \in k$ the λ-eigenspace is a vector subspace of V. The generalized λ-eigenspace is also a vector subspace of V. And both the λ-eigenspace and the generalized one are *stable* under the action of T.

Proof: This is just the linearity of T, hence, of $T - \lambda$. Indeed, for v, w λ-eigenvectors, and for $c \in k$,

$$T(v+w) = Tv + TW = \lambda v + \lambda w = \lambda(v+w) \quad \text{and} \quad T(cv) = c \cdot Tv = c \cdot \lambda v = lam \cdot cv$$

If $(T-\lambda)^m v = 0$ and $(T-\lambda)nw = 0$, let $N = \max(m,n)$. Then

$$(T-\lambda)^N(v+w) = (T-\lambda)^N v + (T-\lambda)^N w = (T-\lambda)^{N-m}(T-\lambda)^m v + (T-\lambda)^{N-n}(T-\lambda)^n w$$

$$= (T-\lambda)^{N-m}0 + (T-\lambda)^{N-n}0 = 0$$

Similarly, generalized eigenspaces are stable under scalar multiplication.

Since the operator T commutes with any polynomial in T, we can compute, for $(T-\lambda)^n v = 0$,

$$(T-\lambda)^n(Tv) = T \cdot (T-\lambda)^n(v) = T(0) = 0$$

which proves the stability. ///

[1.0.2] Proposition: Let $T \in \text{End}_k(V)$ and let $v_1, \ldots, v_m$ be eigenvectors for T, with *distinct* respective eigenvalues $\lambda_1, \ldots, \lambda_m$ in k. Then for scalars c_i

$$c_1 v_1 + \ldots + c_m v_m = 0 \implies \text{all } \ c_i = 0$$

That is, eigenvectors for distinct eigenvalues are linearly independent.

Proof: Suppose that the given relation is the shortest such with all $c_i \neq 0$. Then apply $T - \lambda_1$ to the relation, to obtain

$$0 + (\lambda_2 - \lambda_1)c_2 v_2 \ldots + (\lambda_m - \lambda_1)c_m v_m = 0 \implies \text{all } \ c_i = 0$$

For $i > 1$ the scalars $\lambda_i - \lambda_1$ are not 0, and $(\lambda_i - \lambda_1)v_i$ is again a non-zero λ_i-eigenvector for T. This contradicts the assumption that the relation was the shortest. ///

So far no use was made of finite-dimensionality, and, indeed, all the above arguments are correct without assuming finite-dimensionality. Now, however, we need to assume finite-dimensionality. In particular,

[1.0.3] Proposition: Let V be a finite-dimensional vector space over k. Then

$$\dim_k \text{End}_k(V) = (\dim_k V)^2$$

In particular, $\text{End}_k(V)$ is finite-dimensional.

Proof: An endomorphism T is completely determined by where it sends all the elements of a basis $v_1, \ldots, v_n$ of V, and each v_i can be sent to any vector in V. In particular, let E_{ij} be the endomorphism sending v_i to v_j and sending v_ℓ to 0 for $\ell \neq i$. We claim that these endomorphisms are a k-basis for $\mathrm{End}_k(V)$. First, they span, since any endomorphism T is expressible as

$$T = \sum_{ij} c_{ij} E_{ij}$$

where the $c_{ij} \in k$ are determined by the images of the given basis

$$T(v_i) = \sum_j c_{ij} v_j$$

On the other hand, suppose for some coefficients c_{ij}

$$\sum_{ij} c_{ij}\, E_{ij} = 0 \in \mathrm{End}_k(V)$$

Applying this endomorphism to v_i gives

$$\sum_j c_{ij} v_j = 0 \in V$$

Since the v_j are linearly independent, this implies that all c_{ij} are 0. Thus, the E_{ij} are a basis for the space of endomorphisms, and we have the dimension count. ///

For V finite-dimensional, the homomorphism

$$k[x] \longrightarrow k[T] \subset \mathrm{End}_k(V) \qquad \text{by} \quad x \longrightarrow T$$

from the polynomial ring $k[x]$ to the ring $k[T]$ of polynomials in T must have a non-trivial kernel, since $k[x]$ is infinite-dimensional and $k[T]$ is finite-dimensional. The **minimal polynomial** $f(x) \in k[x]$ of T is the (unique) monic generator of that kernel.

[1.0.4] Proposition: The eigenvalues of a k-linear endomorphism T are exactly the zeros of its minimal polynomial. [1]

Proof: Let $f(x)$ be the minimal polynomial. First, suppose that $x - \lambda$ divides $f(x)$ for some $\lambda \in k$, and put $g(x) = f(x)/(x - \lambda)$. Since $g(x)$ is not divisible by the minimal polynomial, there is $v \in V$ such that $g(T)v \neq 0$. Then

$$(T - \lambda) \cdot g(T)v = f(T) \cdot v = 0$$

so $g(T)v$ is a (non-zero) λ-eigenvector of T. On the other hand, suppose that λ is an eigenvalue, and let v be a non-zero λ-eigenvector for T. If $x - \lambda$ failed to divide $f(x)$, then the *gcd* of $x - \lambda$ and $f(x)$ is 1, and there are polynomials $a(x)$ and $b(x)$ such that

$$1 = a \cdot (x - \lambda) + b \cdot f$$

Mapping $x \longrightarrow T$ gives

$$\mathrm{id}_V = a(T)(T - \lambda) + 0$$

Applying this to v gives

$$v = a(T)(T - \lambda)(v) = a(T) \cdot 0 = 0$$

which contradicts $v \neq 0$. ///

[1] This does not presume that k is algebraically closed.

[1.0.5] **Corollary:** Let k be algebraically closed, and V a finite-dimensional vector space over k. Then there is at least one eigenvalue and (non-zero) eigenvector for any $T \in \mathrm{End}_k(V)$.

Proof: The minimal polynomial has at least one linear factor over an algebraically closed field, so by the previous proposition has at least one eigenvector. ///

[1.0.6] **Remark:** The Cayley-Hamilton theorem[2] is often invoked to deduce the existence of at least one eigenvector, but the last corollary shows that this is not necessary.

2. *Diagonalizability, semi-simplicity*

A linear operator $T \in \mathrm{End}_k(V)$ on a finite-dimensional vector space V over a field k is **diagonalizable**[3] if V has a basis consisting of eigenvectors of T. Equivalently, T may be said to be **semi-simple**, or sometimes V itself, as a $k[T]$ or $k[x]$ module, is said to be semi-simple.

Diagonalizable operators are good, because their effect on arbitrary vectors can be very clearly described as a superposition of scalar multiplications in an obvious manner, namely, letting $v_1, \ldots, v_n$ be eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_n$, if we manage to express a given vector v as a linear combination[4]

$$v = c_1 v_1 + \ldots + c_n v_n$$

of the eigenvectors v_i, with $c_i \in k$, then we can completely describe the effect of T, or even iterates T^ℓ, on v, by

$$T^\ell v = \lambda_1^\ell \cdot c_1 v_1 + \ldots + \lambda_n^\ell \cdot c_n v_n$$

[2.0.1] **Remark:** Even over an algebraically closed field k, an endomorphism T of a finite-dimensional vector space may fail to be diagonalizable by having non-trivial Jordan blocks, meaning that some one of its *elementary divisors* has a *repeated factor*. When k is not necessarily algebraically closed, T may fail to be diagonalizable by having one (hence, at least two) of the zeros of its minimal polynomial lie in a proper field extension of k. For not finite-dimensional V, there are further ways that an endomorphism may fail to be diagonalizable. For example, on the space V of two-sided sequences $a = (\ldots, a_{-1}, a_0, a_1, \ldots)$ with entries in k, the operator T given by

$$i^{th} \text{ component } (Ta)_i \text{ of } Ta = (i-1)^{th} \text{ component } a_i \text{ of } a$$

[2.0.2] **Proposition:** An operator $T \in \mathrm{End}_k(V)$ with V finite-dimensional over the field k is diagonalizable if and only if the minimal polynomial $f(x)$ of T factors into linear factors in $k[x]$ and has no repeated factors. Further, letting V_λ be the λ-eigenspace, diagonalizability is equivalent to

$$V = \sum_{\text{eigenvalues } \lambda} V_\lambda$$

[2] The Cayley-Hamilton theorem, which we will prove later, asserts that the minimal polynomial of an endomorphism T divides the **characteristic polynomial** $\det(T - x \cdot \mathrm{id}_V)$ of T, where det is *determinant*. But this invocation is unnecessary and misleading. Further, it is easy to give false proofs of this result. Indeed, it seems that Cayley and Hamilton only proved the two-dimensional and perhaps three-dimensional cases.

[3] Of course, in coordinates, diagonalizability means that a matrix M giving the endomorphism T can be literally diagonalized by conjugating it by some invertible A, giving diagonal AMA^{-1}. This conjugation amounts to changing coordinates.

[4] The computational problem of expressing a given vector as a linear combination of eigenvectors is not trivial, but is reasonably addressed via *Gaussian elimination.*

Proof: Suppose that f factors into linear factors

$$f(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_n)$$

in $k[x]$ and no factor is repeated. We already saw, above, that the zeros of the minimal polynomial are exactly the eigenvalues, whether or not the polynomial factors into linear factors. What remains is to show that there is a basis of eigenvectors if $f(x)$ factors completely into linear factors, and conversely.

First, suppose that there is a basis $v_1, \ldots, v_n$ of eigenvectors, with eigenvalues $\lambda_1, \ldots, \lambda_n$. Let Λ be the *set*[5] of eigenvalues, specifically *not* attempting to count repeated eigenvalues more than once. Again, we already know that all these eigenvalues do occur among the zeros of the minimal polynomial (*not* counting multiplicities!), and that all zeros of the minimal polynomial are eigenvalues. Let

$$g(x) = \prod_{\lambda \in \Lambda} (x - \lambda)$$

Since every eigenvalue is a zero of $f(x)$, $g(x)$ divides $f(x)$. And $g(T)$ annihilates every eigenvector, and since the eigenvectors span V the endomorphism $g(T)$ is 0. Thus, by definition of the minimal polynomial, $f(x)$ divides $g(x)$. They are both monic, so are equal.

Conversely, suppose that the minimal polynomial $f(x)$ factors as

$$f(x) = (x - \lambda_1) \ldots (x - \lambda_n)$$

with *distinct* λ_i. Again, we have already shown that each λ_i is an eigenvalue. Let V_λ be the λ-eigenspace. Let $\{v_{\lambda,1}, \ldots, v_{lam,d_\lambda}\}$ be a basis for V_λ. We claim that the union

$$\{v_{\lambda,i} : \lambda \text{ an eigenvalue }, \ 1 \le i \le d_\lambda\}$$

of bases for all the (non-trivial) eigenspaces V_λ is a basis for V. We have seen that eigenvectors for *distinct* eigenvalues are linearly independent, so we need only prove

$$\sum_\lambda V_\lambda = V$$

where the sum is over (distinct) eigenvalues. Let $f_\lambda(x) = f(x)/(x - \lambda)$. Since each linear factor occurs only once in f, the *gcd* of the collection of $f_\lambda(x)$ in $k[x]$ is 1. Therefore, there are polynomials $a_\lambda(x)$ such that

$$1 = \gcd(\{f_\lambda : \lambda \text{ an eigenvector}\}) = \sum_\lambda a_\lambda(x) \cdot f_\lambda(x)$$

Then for any $v \in V$

$$v = \mathrm{id}_V(v) = \sum_\lambda a_\lambda(T) \cdot f_\lambda(T)(v)$$

Since

$$(T - \lambda) \cdot f_\lambda(T) = f(T) = 0 \in \mathrm{End}_k(V)$$

[5] Strictly speaking, a set cannot possibly keep track of repeat occurrences, since $\{a, a, b\} = \{a, b\}$, and so on. However, in practice, the notion of set often is corrupted to mean to keep track of repeats. More correctly, a notion of *set* enhanced to keep track of number of repeats is a *multi-set.* Precisely, a **mult-set** M is a set S with a non-negative integer-valued function m on S, where the intent is that $m(s)$ (for $s \in S$) is the number of times s occurs in M, and is called the **multiplicity** of s in M. The question of whether or not the multiplicity can be 0 is a matter of convention and/or taste.

for each eigenvalue λ

$$f_\lambda(T)(V) \subset V_\lambda$$

Thus, in the expression

$$v = \mathrm{id}_V(v) = \sum_\lambda a_\lambda(T) \cdot f_\lambda(T)(v)$$

each $f_\lambda(T)(v)$ is in V_λ. Further, since T and any polynomial in T stabilizes each eigenspace, $a_\lambda(T) f_\lambda(T)(v)$ is in V_λ. Thus, this sum exhibits an arbitrary v as a sum of elements of the eigenspaces, so these eigenspaces do span the whole space.

Finally, suppose that

$$V = \sum_{\text{eigenvalues } \lambda} V_\lambda$$

Then $\prod_\lambda (T - \lambda)$ (product over distinct λ) annihilates the whole space V, so the minimal polynomial of T factors into distinct linear factors. ///

An endomorphism P is a **projector** or **projection** if it is *idempotent*, that is, if

$$P^2 = P$$

The **complementary** or **dual** idempotent is

$$1 - P = \mathrm{id}_V - P$$

Note that

$$(1-P)P = P(1-P) = P - P^2 = 0 \in \mathrm{End}_l(V)$$

Two idempotents P, Q are **orthogonal** if

$$PQ = QP = 0 \in \mathrm{End}_k(V)$$

If we have in mind an endomorphism T, we will usually care only about projectors P *commuting* with T, that is, with $PT = TP$.

[2.0.3] Proposition: Let T be a k-linear operator on a finite-dimensional k-vectorspace V. Let λ be an eigenvalue of T, with eigenspace V_λ, and suppose that the factor $x - \lambda$ occurs with multiplicity one in the minimal polynomial $f(x)$ of T. Then there is a polynomial $a(x)$ such that $a(T)$ is a *projector commuting with* T, and is the identity map on the λ-eigenspace.

Proof: Let $g(x) = f(x)/(x-\lambda)$. The multiplicity assumption assures us that $x - \lambda$ and $g(x)$ are relatively prime, so there are $a(x)$ and $b(x)$ such that

$$1 = a(x)g(x) + b(x)(x - \lambda)$$

or

$$1 - b(x)(x-\lambda) = a(x)g(x)$$

As in the previous proof, $(x-\lambda)g(x) = f(x)$, so $(T-\lambda)g(T) = 0$, and $g(T)(V) \subset V_\lambda$. And, further, because T and polynomials in T stabilize eigenspaces, $a(T)g(T)(V) \subset V_\lambda$. And

$$[a(T)g(T)]^2 = a(T)g(T) \cdot [1 - b(T)(T-\lambda)] = a(T)g(T) - 0 = a(T)g(T)$$

since $g(T)(T-\lambda) = f(T) = 0$. That is,

$$P = a(T)g(T) \cdot$$

is the desired projector to the λ-eigenspace. ///

[2.0.4] **Remark:** The condition that the projector commute with T is non-trivial, and without it there are many projectors that will not be what we want.

3. *Commuting endomorphisms* $ST = TS$

Two endomorphisms $S, T \in \mathrm{End}_k(V)$ are said to **commute** (with each other) if

$$ST = TS$$

This hypothesis allows us to reach some worthwhile conclusions about eigenvectors of the two separately, and jointly. Operators which do not commute are much more complicated to consider from the viewpoint of eigenvectors.[6]

[3.0.1] **Proposition:** Let S, T be commuting endomorphisms of V. Then S stabilizes every eigenspace of T.

Proof: Let v be a λ-eigenvector of T. Then

$$T(Sv) = (TS)v = (ST)v = S(Tv) = S(\lambda v) = \lambda \cdot Sv$$

as desired. ///

[3.0.2] **Proposition:** Commuting *diagonalizable* endomorphisms S and T on V are *simultaneously diagonalizable*, in the sense that there is a basis consisting of vectors which are simultaneously eigenvectors for *both* S and T.

Proof: Since T is diagonalizable, from above V decomposes as

$$V = \sum_{\text{eigenvalues } \lambda} V_\lambda$$

where V_λ is the λ-eigenspace of T on V. From the previous proposition, S stabilizes each V_λ.

Let's (re) prove that for S diagonalizable on a vector space V, that S is diagonalizable on any S-stable subspace W. Let $g(x)$ be the minimal polynomial of S on V. Since W is S-stable, it makes sense to speak of the minimal polynomial $h(x)$ of S on W. Since $g(S)$ annihilates V, it certainly annihilates W. Thus, $g(x)$ is a polynomial multiple of $h(x)$, since the latter is the unique monic generator for the ideal of polynomials $P(x)$ such that $P(S)(W) = 0$. We proved in the previous section that the diagonalizability of S on V implies that $g(x)$ factors into linear factors in $k[x]$ and no factor is repeated. Since $h(x)$ divides $g(x)$, the same is true of $h(x)$. We saw in the last section that this implies that S on W is diagonalizable.

In particular, V_λ has a basis of eigenvectors for S. These are all λ-eigenvectors for T, so are indeed *simultaneous* eigenvectors for the two endomorphisms. ///

4. *Inner product spaces*

Now take the field k to be either $\mathbb{R}$ or $\mathbb{C}$. We use the **positivity property** of $\mathbb{R}$ that for $r_1, \ldots, r_n \in \mathbb{R}$

$$r_1^2 + \ldots + r_n^2 = 0 \implies \text{all } r_i = 0$$

[6] Indeed, to study non-commutative collections of operators the notion of *eigenvector* becomes much less relevant. Instead, a more complicated (and/but more interesting) notion of *irreducible subspace* is the proper generalization.

The **norm-squared** of a complex number $\alpha = a + bi$ (with $a, b \in \mathbb{R}$) is

$$|\alpha|^2 = \alpha \cdot \overline{\alpha} = a^2 + b^2$$

where $\overline{a+bi} = a - bi$ is the usual **complex conjugative**. The positivity property in $\mathbb{R}$ thus implies an analogous one for $\alpha_1, \ldots, \alpha_n$, namely

$$|\alpha_1|^2 + \ldots + |\alpha_n|^2 = 0 \quad \Longrightarrow \quad \text{all } \alpha_i = 0$$

[4.0.1] Remark: In the following, for scalars $k = \mathbb{C}$ we will need to refer to the complex conjugation on it. But when k is $\mathbb{R}$ the conjugation is trivial. To include both cases at once we will systematically refer to *conjugation*, with the reasonable convention that for $k = \mathbb{R}$ this is the do-nothing operation.

Given a k-vectorspace V, an **inner product** or **scalar product** or **dot product** or **hermitian product** (the latter especially if the set k of scalars is $\mathbb{C}$) is a k-valued function

$$\langle , \rangle : V \times V \longrightarrow k$$

written

$$v \times w \longrightarrow \langle v, w \rangle$$

which meets several conditions. First, a mild condition that $\langle , \rangle$ be k-linear in the first argument and k-conjugate-linear in the second, meaning that $\langle , \rangle$ is *additive* in both arguments:

$$\langle v + v', w + w' \rangle = \langle v, w \rangle + \langle v', w \rangle + \langle v, w' \rangle + \langle v', w' \rangle$$

and scalars behave as

$$\langle \alpha v, \beta w \rangle = \alpha \, \overline{\beta} \, \langle v, w \rangle$$

The inner product is **hermitian** in the sense that

$$\langle v, w \rangle = \overline{\langle w, v \rangle}$$

Thus, for ground field k either $\mathbb{R}$ or $\mathbb{C}$,

$$\langle v, v \rangle = \overline{\langle v, v \rangle}$$

so $\langle v, v \rangle \in \mathbb{R}$.

The most serious condition on $\langle , \rangle$ is **positive-definiteness**, which is that

$$\langle v, v \rangle \geq 0 \quad \text{with equality only for } v = 0$$

Two vectors v, w are **orthogonal** or **perpendicular** if

$$\langle v, w \rangle = 0$$

We may write $v \perp w$ for the latter condition. There is an associated **norm**

$$|v| = \langle v, v \rangle^{1/2}$$

and **metric**

$$d(v, w) = |v - w|$$

A vector space basis $e_1, e_2, \ldots, e_n$ of V is an **orthonormal basis** for V if

$$\langle e_i, e_j \rangle = \begin{cases} 1 & (\text{for } i = j) \\ 1 & (\text{for } i = j) \\ 0 & (\text{for } i \neq j) \end{cases}$$

[4.0.2] Proposition: *(Gram-Schmidt process)* Given a basis $v_1, v_2, \ldots, v_n$ of a finite-dimensional inner product space V, let

$$\begin{array}{rclcrcl} & & & & e_1 & = & \frac{v_1}{|v_1|} \\ v_2' & = & v_2 - \langle v_2, e_1\rangle e_1 & \text{and} & e_2 & = & \frac{v_2'}{|v_2'|} \\ v_3' & = & v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2 & \text{and} & e_3 & = & \frac{v_3'}{|v_3'|} \\ & \ldots & & & & & \\ v_i' & = & v_i - \sum_{j<i} \langle v_i, e_j\rangle e_j & \text{and} & e_i & = & \frac{v_i'}{|v_i'|} \\ & \ldots & & & & & \end{array}$$

Then $e_1, \ldots, e_n$ is an orthonormal basis for V.

[4.0.3] Remark: One could also give a more existential proof that orthonormal bases exist, but the conversion of arbitrary basis to an orthonormal one is of additional interest.

Proof: Use induction. Note that for any vector e of length 1

$$\langle v - \langle v, e\rangle e, e\rangle = \langle v, e\rangle - \langle v, e\rangle\langle e, e\rangle = \langle v, e\rangle - \langle v, e\rangle \cdot 1 = 0$$

Thus, for $\ell < i$,

$$\langle v_i', e_\ell\rangle = \langle v_i - \sum_{j<i} \langle v_i, e_j\rangle e_j, e_\ell\rangle = \langle v_i, e_\ell\rangle - \langle\langle v_i, e_\ell\rangle e_\ell, e_\ell\rangle - \sum_{j<i,\, j\neq\ell} \langle\langle v_i, e_j\rangle e_j, e_\ell\rangle$$

$$= \langle v_i, e_\ell\rangle - \langle v_i, e_\ell\rangle\langle e_\ell, e_\ell\rangle - \sum_{j<i,\, j\neq\ell} \langle v_i, e_j\rangle\langle e_j, e_\ell\rangle = \langle v_i, e_\ell\rangle - \langle v_i, e_\ell\rangle - 0 = 0$$

since the e_j's are (by induction) mutually orthogonal and have length 1. One reasonable worry is that v_i' is 0. But by induction $e_1, e_2, \ldots, e_{i-1}$ is a basis for the subspace of V for which $v_1, \ldots, v_{i-1}$ is a basis. Thus, since v_i is linearly independent of $v_1, \ldots, v_{i-1}$ it is also independent of $e_1, \ldots, e_{i-1}$, so the expression

$$v_i' = v_i + (\text{linear combination of } e_1, \ldots, e_{i-1})$$

cannot give 0. Further, that expression gives the induction step proving that the span of $e_1, \ldots, e_i$ is the same as that of $v_1, \ldots, v_i$. ///

Let W be a subspace of a finite-dimensional k-vectorspace (k is $\mathbb{R}$ or $\mathbb{C}$) with a (positive-definite) inner product $\langle, \rangle$. The **orthogonal complement** $W^\top$ is

$$W^\top = \{v \in V : \langle v, w\rangle = 0 \text{ for all } w \in W\}$$

It is easy to check that the orthogonal complement is a vector subspace.

[4.0.4] Theorem: In finite-dimensional vector spaces V, for subspaces W[7]

$$W^{\perp^\perp} = W$$

In particular, for any W

$$\dim_k W + \dim_k W^\perp = \dim_k V$$

Indeed,

$$V = W \oplus W^\perp$$

[7] In infinite-dimensional inner-product spaces, the orthogonal complement of the orthogonal complement is the topological *closure* of the original subspace.

There is a unique projector P which is an **orthogonal projector** to W in the sense that on P is the identity on W and is 0 on $W^\perp$.

Proof: First, we verify some relatively easy parts. For $v \in W \cap W^\perp$ we have $0 = \langle v, v \rangle$, so $v = 0$ by the positive-definiteness. Next, for $w \in W$ and $v \in W^\perp$,

$$0 = \langle w, v \rangle = \overline{\langle v, w \rangle} = \overline{0}$$

which proves this inclusion $W \subset W^{\perp\perp}$.

Next, suppose that for a given $v \in V$ there were two expressions

$$v = w + w' = u + u'$$

with $w, u \in W$ and $w', u' \in W^\perp$. Then

$$W \ni w - u = u' - w' \in W^\perp$$

Since $W \cap W^\perp = 0$, it must be that $w = u$ and $w' = u'$, which gives the uniqueness of such an expression (assuming existence).

Let $e_1, \ldots, e_m$ be an orthogonal basis for W. Given $v \in V$, let

$$x = \sum_{1 \le i \le m} \langle v, e_i \rangle \, e_i$$

and

$$y = v - x$$

Since it is a linear combination of the e_i, certainly $x \in W$. By design, $y \in W^\perp$, since for any e_ℓ

$$\langle y, w \rangle = \langle v - \sum_{1 \le i \le m} \langle v, e_i \rangle \, e_i, \; e_\ell \rangle = \langle v, e_\ell \rangle - \sum_{1 \le i \le m} \langle v, e_i \rangle \, \langle e_i, e_\ell \rangle = \langle v, e_\ell \rangle - \langle v, e_\ell \rangle$$

since the e_i are an orthonormal basis for W. This expresses

$$v = x + y$$

as a linear combination of elements of W and $W^\perp$.

Since the map $v \longrightarrow x$ is expressible in terms of the inner product, as just above, this is the desired projector to W. By the uniqueness of the decomposition into W and $W^\perp$ components, the projector is orthogonal, as desired. ///

[4.0.5] Corollary: [8] Suppose that a finite-dimensional vector space V has an inner product $\langle , \rangle$. To every k-linear map $L : V \longrightarrow k$ is attached a unique $w \in V$ such that for all $v \in V$

$$Lv = \langle v, w \rangle$$

[4.0.6] Remark: The k-linear maps of a k-vectorspace V to k itself are called **linear functionals** on V.

[8] This is a very simple case of the Riesz-Fischer theorem, which asserts the analogue for *Hilbert* spaces, which are the proper infinite-dimensional version of inner-product spaces. In particular, Hilbert spaces are required, in addition to the properties mentioned here, to be *complete* with respect to the metric $d(x, y) = |x - y|$ coming from the inner product. This completeness is automatic for finite-dimensional inner product spaces.

Proof: If L is the 0 map, just take $w = 0$. Otherwise, since

$$\dim_k \ker L = \dim_k V - \dim_k \operatorname{Im} L = \dim_k V - \dim_k k = \dim_k V - 1$$

Take a vector e of length 1 in the orthogonal complement[9] $(\ker L)^\perp$. For arbitrary $v \in V$

$$v - \langle v, e\rangle e \in \ker L$$

Thus,

$$L(v) = L(v - \langle v, e\rangle\, e) + L(\langle v, e\rangle\, e) = 0 + \langle v, e\rangle L(e) = \langle v, \overline{L(e)}e\rangle$$

That is, $w = \overline{L(e)}e$ is the desired element of V. ///

The **adjoint** T^* of $T \in \operatorname{End}_k(V)$ with respect to an inner product $\langle,\rangle$ is another linear operator in $\operatorname{End}_k(V)$ such that, for all $v, w \in V$,

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

[4.0.7] Proposition: Adjoint operators (on finite-dimensional inner product spaces) exist and are unique.

Proof: Let T be a linear endomorphism of V. Given $x \in V$, the map $v \longrightarrow \langle Tv, x\rangle$ is a linear map to k. Thus, by the previous corollary, there is a unique $y \in V$ such that for all $v \in V$

$$\langle Tv, x\rangle = \langle v, y\rangle$$

We want to define $T^*x = y$. This is well-defined as a function, but we need to prove linearity, which, happily, is not difficult. Indeed, let $x, x' \in V$ and let y, y' be attached to them as just above. Then

$$\langle Tv, x + x'\rangle = \langle Tv, x\rangle + \langle Tv, x\rangle = \langle v, y\rangle + \langle v, y'\rangle = \langle v, y + y'\rangle$$

proving the additivity $T^*(x + x') = T * x + T^*x'$. Similarly, for $c \in k$,

$$\langle Tv, cx\rangle = \bar{c}\langle Tv, x\rangle = \bar{c}\langle v, y\rangle = \langle v, cy\rangle$$

proving the linearity of T^*. ///

Note that the direct computation

$$\langle T * v, w\rangle = \overline{\langle w, T * v\rangle} = \overline{\langle Tw, v\rangle} = \langle v, Tw, v\rangle$$

shows that, unsurprisingly,

$$(T^*)^* = T$$

A linear operator T on an inner product space V is **normal**[10] if it commutes with its adjoint, that is, if

$$TT^* = T^*T$$

An operator T is **self-adjoint** or **hermitian** if it is equal to its adjoint, that is, if

$$T = T^*$$

[9] Knowing that the orthogonal complement exists is a crucial point, and that fact contains more information than is immediately apparent.

[10] Yet another oh-so-standard but unhelpful use of this adjective.

An operator T on an inner product space V is **unitary** if[11]

$$T^*T = \mathrm{id}_V$$

Since we are discussing finite-dimensional V, this implies that the kernel of T is trivial, and thus T is invertible, since (as we saw much earlier)

$$\dim\ker T + \dim \mathrm{Im} T = \dim V$$

[4.0.8] **Proposition:** Eigenvalues of self-adjoint operators T on an inner product space V are *real.*

Proof: Let v be a (non-zero) eigenvector for T, with eigenvalue λ. Then

$$\lambda\langle v,v\rangle = \langle \lambda v,v\rangle = \langle Tv,v\rangle = \langle v,T^*v\rangle = \langle v,Tv\rangle = \langle v,\lambda v\rangle = \overline{\lambda}\langle v,v\rangle$$

Since $\langle v,v\rangle \neq 0$, this implies that $\overline{\lambda} = \lambda$. ///

5. *Projections without coordinates*

There is another construction of orthogonal projections and orthogonal complements which is less coordinate-dependent, and which applies to infinite-dimensional [12] inner-product spaces as well. Specifically, using the metric

$$d(x,y) = |x-y| = \langle x-y, x-y\rangle^{1/2}$$

the **orthogonal projection** of a vector x to the subspace W is the vector in W closest to x.

To prove this, first observe the **polarization identity**

$$|x+y|^2 + |x-y|^2 = |x|^2 + \langle x,y\rangle + \langle y,x\rangle + |y|^2 + |x|^2 - \langle x,y\rangle - \langle y,x\rangle + |y|^2 = 2|x|^2 + 2|y|^2$$

Fix x not in W, and let u, v be in W such that $|x-u|^2$ and $|x-v|^2$ are within $\varepsilon > 0$ of the infimum μ of all values $|x-w|^2$ for $w \in W$. Then an application of the previous identity gives

$$|(x-u)+(x-v)|^2 + |(x-u)-(x-v)|^2 = 2|x-u|^2 + 2|x-v|^2$$

so

$$|u-v|^2 = 2|x-u|^2 + 2|x-v|^2 - |(x-u)+(x-v)|^2$$

The further small trick is to notice that

$$(x-u)+(x-v) = 2\cdot(x - \frac{u+v}{2})$$

which is again of the form $x - w'$ for $w' \in W$. Thus,

$$|u-v|^2 = 2|x-u|^2 + 2|x-v|^2 - 4|x-\frac{u+v}{2}|^2 < 2(\mu+\varepsilon) + 2(\mu+\varepsilon) - 4\mu = 4\varepsilon$$

[11] For infinite-dimensional spaces this definition of *unitary* is insufficient. Invertibility must be explicitly required, one way or another.

[12] Precisely, this argument applies to arbitrary inner product spaces that are *complete* in the metric sense, namely, that Cauchy sequences converge in the metric naturally attached to the inner product, namely $d(x,y) = |x-y| = \langle x-y, x-y\rangle^{1/2}$.

That is, we can make a Cauchy sequence from the u, v.

Granting that Cauchy sequences converge, this proves *existence* of a closest point of W to x, as well as the *uniqueness* of the closest point. ///

From this viewpoint, the **orthogonal complement** $W^\perp$ to W can be defined to be the collection of vectors x in V such that the orthogonal projection of x to W is 0.

6. *Unitary operators*

It is worthwhile to look at different ways of characterizing and constructing *unitary* operators on a finite-dimensional complex vector space V with a hermitian inner product $\langle , \rangle$. These equivalent conditions are easy to verify once stated, but it would be unfortunate to overlook them, so we make them explicit. Again, the *definition* of the unitariness of $T : V \longrightarrow V$ for finite-dimensional[13] V is that $T^*T = \mathrm{id}_V$.

[6.0.1] **Proposition:** For V finite-dimensional[14] $T \in \mathrm{End}_{\mathbb{C}}(V)$ is unitary if and only if $TT^* = \mathrm{id}_V$. Unitary operators on finite-dimensional spaces are necessarily invertible.

Proof: The condition $T^*T = \mathrm{id}_V$ implies that T is injective (since it has a left inverse), and since V is finite-dimensional T is also surjective, so is an isomorphism. Thus, its left inverse T^* is also its right inverse, by uniqueness of inverses. ///

[6.0.2] **Proposition:** For V finite-dimensional with hermitian inner product $\langle , \rangle$ an operator $T \in \mathrm{End}_{\mathbb{C}}(V)$ is unitary if and only if

$$\langle Tu, Tv \rangle = \langle u, v \rangle$$

for all $u, v \in V$.

Proof: If $T^*T = \mathrm{id}_V$, then by definition of adjoint

$$\langle Tu, Tv \rangle = \langle T^*Tu, v \rangle = \langle \mathrm{id}_V u, v \rangle = \langle u, v \rangle$$

On the other hand, if

$$\langle Tu, Tv \rangle = \langle u, v \rangle$$

then

$$0 = \langle T^*Tu, v \rangle - \langle u, v \rangle = \langle (T^*T - \mathrm{id}_V)u, v \rangle$$

Take $v = (T^*T - \mathrm{id}_V)u$ and invoke the positivity of $\langle , \rangle$ to conclude that $(T^*T - \mathrm{id}_V)u = 0$ for all u. Thus, as an endomorphism, $T^*T - \mathrm{id}_V = 0$, and T is unitary. ///

[6.0.3] **Proposition:** For a unitary operator $T \in \mathrm{End}_{\mathbb{C}}(V)$ on a finite-dimensional V with hermitian inner product $\langle , \rangle$, and for given orthonormal basis $\{f_i\}$ for V, the set $\{Tf_i\}$ is also an orthonormal basis. Conversely, given two ordered orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ for V, the uniquely determined endomorphism T such that $Te_i = f_i$ is unitary.

Proof: The first part is immediate. For an orthonormal basis $\{e_i\}$ and unitary T,

$$\langle Te_i, Te_j \rangle = \langle e_i, e_j \rangle$$

[13] For infinite-dimensional V one must also explicitly require that T be invertible to have the best version of unitariness. In the finite-dimensional case the first proposition incidentally shows that invertibility is automatic.

[14] Without finite-dimensionality this assertion is generally false.

so the images Te_i make up an orthonormal basis.

The other part is still easy, but requires a small computation whose idea is important. First, since e_i form a basis, there is a unique linear endomorphism T sending e_i to any particular chosen ordered list of targets. To prove the unitariness of this T we use the criterion of the previous proposition. Let $u = \sum_i a_i e_i$ and $v = \sum_j b_j e_j$ with a_i and b_j in $\mathbb{C}$. Then, on one hand,

$$\langle Tu, Tv \rangle = \sum_{ij} a_i \bar{b}_j \, \langle Te_i, Te_j \rangle = \sum_i a_i \bar{b}_i$$

by the hermitian-ness of $\langle , \rangle$ and by the linearity of T. On the other hand, a very similar computation gives

$$\langle u, v \rangle = \sum_{ij} a_i \bar{b}_j \, \langle Te_i, Te_j \rangle = \sum_i a_i \bar{b}_i$$

Thus, T preserves inner products, so is unitary. ///

7. *Spectral theorems*

The spectral theorem[15] for *normal* operators subsumes the spectral theorem for *self-adjoint* operators, but the proof in the self-adjoint case is so easy to understand that we give this proof separately. Further, many of the applications to matrices use only the self-adjoint case, so understanding this is sufficient for many purposes.

[7.0.1] **Theorem:** Let T be a self-adjoint operator on a finite-dimensional complex vector space V with a (hermitian) inner product $\langle , \rangle$. Then there is an orthonormal basis $\{e_i\}$ for V consisting of eigenvectors for T.

Proof: To prove the theorem, we need

[7.0.2] **Proposition:** Let W be a T-stable subspace of V, with $T = T^*$. Then the orthogonal complement $W^\perp$ is also T-stable.

Proof: *(of proposition)* Let $v \in W^\perp$, and $w \in W$. Then

$$\langle Tv, w \rangle = \langle v, T^* w \rangle = \langle v, Tw \rangle = 0$$

since $Tw \in W$. ///

To prove the theorem, we do an induction on the dimension of V. Let $v \neq 0$ be any vector of length 1 which is an eigenvector for T. We know that T has eigenvectors simply because $\mathbb{C}$ is algebraically closed (so the minimal polynomial of T factors into linear factors) and V is finite-dimensional. Thus, $\mathbb{C} \cdot v$ is T-stable, and, by the proposition just proved, the orthogonal complement $(\mathbb{C} \cdot v)^\perp$ is also T-stable. With the restriction of the inner product to $(\mathbb{C} \cdot v)^\perp$ the restriction of T is *still* self-adjoint, so by induction on dimension we're done. ///

Now we give the more general, and somewhat more complicated, argument for normal operators. This does include the previous case, as well as the case of unitary operators.

[15] The use of the word *spectrum* is a reference to wave phenomena, and the idea that a complicated wave is a superposition of simpler ones.

[7.0.3] **Theorem:** Let T be a normal operator on a finite-dimensional complex vector space V with a (hermitian) inner product $\langle , \rangle$. Then there is an orthonormal basis $\{e_i\}$ for V consisting of eigenvectors for T.

Proof: First prove

[7.0.4] **Proposition:** Let T be an operator on V, and W a T-stable subspace. Then the orthogonal complement $W^\perp$ of W is T^*-stable. [16]

Proof: (of proposition) Let $v \in W^\perp$, and $w \in W$. Then

$$\langle T^* v, w\rangle = \langle v, Tw\rangle = 0$$

since $Tw \in W$. ///

The proof of the theorem is by induction on the dimension of V. Let λ be an eigenvalue of T, and V_λ the λ-eigenspace of T on V. The assumption of normality is that T and T^* commute, so, from the general discussion of commuting operators, T^* stabilizes V_λ. Then, by the proposition just proved, $T = T^{**}$ stabilizes $V_\lambda^\perp$. By induction on dimension, we're done. ///

8. *Corollaries of the spectral theorem*

These corollaries do not mention the spectral theorem, so do not hint that it plays a role.

[8.0.1] **Corollary:** Let T be a self-adjoint operator on a finite-dimensional complex vector space V with inner product $\langle , \rangle$. Let $\{e_i\}$ be an orthonormal basis for V. Then there is a unitary operator k on V (that is, $\langle kv, kw\rangle = \langle v, w\rangle$ for all $v, w \in V$) such that

$$\{ke_i\} \quad \text{is an orthonormal basis of } T\text{-eigenvectors}$$

Proof: Let $\{f_i\}$ be an orthonormal basis of T-eigenvectors, whose existence is assured by the spectral theorem. Let k be a linear endomorphism mapping $e_i \longrightarrow f_i$ for all indices i. We claim that k is unitary. Indeed, letting $v = \sum_i a_i e_i$ and $w = \sum_j b_j e_j$,

$$\langle kv, kw\rangle = \sum_{ij} a_i \bar{b}_j \, \langle ke_i, ke_j\rangle = \sum_{ij} a_i \bar{b}_j \, \langle f_i, f_j\rangle = \sum_{ij} a_i \bar{b}_j \langle e_i, e_j\rangle = \langle v, w\rangle$$

This is the unitariness. ///

A self-adjoint operator T on a finite-dimensional complex vector space V with hermitian inner product is **positive definite** if

$$\langle Tv, v\rangle \geq 0 \quad \text{with equality only for } \; v = 0$$

The operator T is **positive semi-definite** if

$$\langle Tv, v\rangle \geq 0$$

(that is, equality may occur for non-zero vectors v).

[8.0.2] **Proposition:** The eigenvalues of a positive definite operator T are positive real numbers. When T is merely positive semi-definite, the eigenvalues are non-negative.

[16] Indeed, this is the natural extension of the analogous proposition in the theorem for hermitian operators.

Proof: We already showed that the eigenvalues of a self-adjoint operator are real. Let v be a non-zero λ-eigenvector for T. Then

$$\lambda\langle v, v\rangle = \langle Tv, v\rangle > 0$$

by the positive definiteness. Since $\langle v, v\rangle > 0$, necessarily $\lambda > 0$. When T is merely semi-definite, we get only $\lambda \geq 0$ by this argument. ///

[8.0.3] **Corollary:** Let $T = T^*$ be positive semi-definite. Then T has a positive semi-definite square root S, that is, S is self-adjoint, positive semi-definite, and

$$S^2 = T$$

If T is positive definite, then S is positive definite.

Proof: Invoking the spectral theorem, there is an orthonormal basis $\{e_i\}$ for V consisting of eigenvectors, with respective eigenvalues $\lambda_i \geq 0$. Define an operator S by

$$Se_i = \sqrt{\lambda_i} \cdot e_i$$

Clearly S has the same eigenvectors as T, with eigenvalues the non-negative real square roots of those of T, and the square of this operator is T. We check directly that it is self-adjoint: let $v = \sum_i a_i e_i$ and $w = \sum_i b_i e_i$ and compute

$$\langle S^* v, w\rangle = \langle v, Sw\rangle = \sum_{ij} a_i \overline{b_j} \langle e_i, e_j\rangle = \sum_{ij} a_i \overline{b_j} \sqrt{\lambda_j} \langle e_i, e_j\rangle = \sum_i a_i \overline{b_i} \sqrt{\lambda_i} \langle e_i, e_i\rangle$$

by orthonormality and the real-ness of $\sqrt{\lambda_i}$. Going backwards, this is

$$\sum_{ij} a_i \overline{b_j} \langle \sqrt{\lambda_i} e_i, e_j\rangle = \langle Sv, w\rangle$$

Since the adjoint is unique, $S = S^*$. ///

The **standard (hermitian) inner product** on $\mathbb{C}^n$ is

$$\langle (v_1, \ldots, v_n),\ (w_1, \ldots, w_n)\rangle = \sum_{i=1}^{n} v_i \overline{w_j}$$

In this situation, certainly n-by-n complex matrices give $\mathbb{C}$ linear endomorphisms by left multiplication of column vectors. With this inner product, the adjoint of an endomorphism T is

$$T^* = T - \text{conjugate-transpose}$$

as usual. Indeed, we often write the superscript-star to indicate conjugate-transpose of a matrix, if no other meaning is apparent from context, and say that the matrix T is **hermitian**. Similarly, an n-by-n matrix k is **unitary** if

$$kk^* = 1_n$$

where 1_n is the n-by-n identity matrix. This is readily verified to be equivalent to unitariness with respect to the standard hermitian inner product.

[8.0.4] **Corollary:** Let T be a hermitian matrix. Then there is a unitary matrix k such that

$$k^* T k = \text{ diagonal, with diagonal entries the eigenvalues of } T$$

Proof: Let $\{e_i\}$ be the standard basis for $\mathbb{C}^n$. It is orthonormal with respect to the standard inner product. Let $\{f_i\}$ be an orthonormal basis consisting of T-eigenvectors. From the first corollary of this section, let k be the unitary operator mapping e_i to f_i. Then k^*Tk is diagonal, with diagonal entries the eigenvalues. ///

[8.0.5] **Corollary:** Let T be a positive semi-definite hermitian matrix. Then there is a positive semi-definite hermitian matrix S such that

$$S^2 = T$$

Proof: With respect to the standard inner product T is positive semi-definite self-adjoint, so has such a square root, from above. ///

9. *Worked examples*

[24.1] Let p be the smallest prime dividing the order of a finite group G. Show that a subgroup H of G of index p is necessarily *normal*.

Let G act on cosets gH of H by left multiplication. This gives a homomorphism f of G to the group of permutations of $[G : H] = p$ things. The kernel $\ker f$ certainly lies inside H, since $gH = H$ only for $g \in H$. Thus, $p|[G : \ker f]$. On the other hand,

$$|f(G)| = [G : \ker f] = |G|/|\ker f|$$

and $|f(G)|$ divides the order $p!$ of the symmetric group on p things, by Lagrange. But p is the smallest prime dividing $|G|$, so $f(G)$ can only have order 1 or p. Since p divides the order of $f(G)$ and $|f(G)|$ divides p, we have equality. That is, H is the kernel of f. Every kernel is normal, so H is normal. ///

[24.2] Let $T \in \mathrm{Hom}_k(V)$ for a finite-dimensional k-vectorspace V, with k a field. Let W be a T-stable subspace. Prove that the minimal polynomial of T on W is a divisor of the minimal polynomial of T on V. Define a natural action of T on the quotient V/W, and prove that the minimal polynomial of T on V/W is a divisor of the minimal polynomial of T on V.

Let $f(x)$ be the minimal polynomial of T on V, and $g(x)$ the minimal polynomial of T on W. (We need the T-stability of W for this to make sense at all.) Since $f(T) = 0$ on V, and since the restriction map

$$\mathrm{End}_k(V) \longrightarrow \mathrm{End}_k(W)$$

is a ring homomorphism,

$$(\text{restriction of})f(t) = f(\text{restriction of } T)$$

Thus, $f(T) = 0$ on W. That is, by definition of $g(x)$ and the PID-ness of $k[x]$, $f(x)$ is a multiple of $g(x)$, as desired.

Define $\overline{T}(v + W) = Tv + W$. Since $TW \subset W$, this is well-defined. Note that we cannot assert, and do not need, an *equality* $TW = W$, but only containment. Let $h(x)$ be the minimal polynomial of $\overline{T}$ (on V/W). Any polynomial $p(T)$ stabilizes W, so gives a well-defined map $\overline{p(T)}$ on V/W. Further, since the natural map

$$\mathrm{End}_k(V) \longrightarrow \mathrm{End}_k(V/W)$$

is a ring homomorphism, we have

$$\overline{p(T)}(v + W) = p(T)(v) + W = p(T)(v + W) + W = p(\overline{T})(v + W)$$

Since $f(T) = 0$ on V, $f(\overline{T}) = 0$. By definition of minimal polynomial, $h(x)|f(x)$. ///

[24.3] Let $T \in \mathrm{Hom}_k(V)$ for a finite-dimensional k-vectorspace V, with k a field. Suppose that T is *diagonalizable* on V. Let W be a T-stable subspace of V. Show that T is diagonalizable on W.

Since T is diagonalizable, its minimal polynomial $f(x)$ on V factors into linear factors in $k[x]$ (with zeros exactly the eigenvalues), and no factor is repeated. By the previous example, the minimal polynomial $g(x)$ of T on W divides $f(x)$, so (by unique factorization in $k[x]$) factors into linear factors without repeats. And this implies that T is diagonalizable when restricted to W. ///

[24.4] Let $T \in \mathrm{Hom}_k(V)$ for a finite-dimensional k-vectorspace V, with k a field. Suppose that T is *diagonalizable* on V, with *distinct eigenvalues*. Let $S \in \mathrm{Hom}_k(V)$ commute with T, in the natural sense that $ST = TS$. Show that S is diagonalizable on V.

The hypothesis of *distinct eigenvalues* means that each eigenspace is *one-dimensional.* We have seen that commuting operators stabilize each other's eigenspaces. Thus, S stabilizes each one-dimensional λ-eigenspaces V_λ for T. By the one-dimensionality of V_λ, S is a scalar μ_λ on V_λ. That is, the basis of eigenvectors for T is unavoidably a basis of eigenvectors for S, too, so S is diagonalizable. ///

[24.5] Let $T \in \mathrm{Hom}_k(V)$ for a finite-dimensional k-vectorspace V, with k a field. Suppose that T is *diagonalizable* on V. Show that $k[T]$ contains the projectors to the eigenspaces of T.

Though it is only implicit, we only want projectors P which *commute* with T.

Since T is diagonalizable, its minimal polynomial $f(x)$ factors into linear factors and has no repeated factors. For each eigenvalue λ, let $f_\lambda(x) = f(x)/(x-\lambda)$. The hypothesis that no factor is repeated implies that the *gcd* of all these $f_\lambda(x)$ is 1, so there are polynomials $a_\lambda(x)$ in $k[x]$ such that

$$1 = \sum_\lambda a_\lambda(x)\, f_\lambda(x)$$

For $\mu \neq \lambda$, the product $f_\lambda(x) f_\mu(x)$ picks up all the linear factors in $f(x)$, so

$$f_\lambda(T) f_\mu(T) = 0$$

Then for each eigenvalue μ

$$(a_\mu(T)\, f_\mu(T))^2 = (a_\mu(T)\, f_\mu(T))\,(1 - \sum_{\lambda\neq\mu} a_\lambda(T)\, f_\lambda(T)) = (a_\mu(T)\, f_\mu(T))$$

Thus, $P_\mu = a_\mu(T)\, f_\mu(T)$ has $P_\mu^2 = P_\mu$. Since $f_\lambda(T) f_\mu(T) = 0$ for $\lambda \neq \mu$, we have $P_\mu P_\lambda = 0$ for $\lambda \neq \mu$. Thus, these are projectors to the eigenspaces of T, and, being polynomials in T, commute with T.

For uniqueness, observe that the diagonalizability of T implies that V is the sum of the λ-eigenspaces V_λ of T. We know that any endomorphism (such as a projector) commuting with T stabilizes the eigenspaces of T. Thus, given an eigenvalue λ of T, an endomorphism P commuting with T and such that $P(V) = V_\lambda$ must be 0 on T-eigenspaces V_μ with $\mu \neq \lambda$, since

$$P(V_\mu) \subset V_\mu \cap V_\lambda = 0$$

And when restricted to V_λ the operator P is required to be the identity. Since V is the sum of the eigenspaces and P is determined completely on each one, there is only one such P (for each λ). ///

[24.6] Let V be a complex vector space with a (positive definite) inner product. Show that $T \in \mathrm{Hom}_k(V)$ cannot be a normal operator if it has any non-trivial Jordan block.

The spectral theorem for normal operators asserts, among other things, that normal operators are diagonalizable, in the sense that there is a basis of eigenvectors. We know that this implies that the minimal polynomial has no repeated factors. Presence of a non-trivial Jordan block exactly means that the minimal polynomial *does* have a repeated factor, so this cannot happen for normal operators. ///

[24.7] Show that a positive-definite hermitian n-by-n matrix A has a unique positive-definite square root B (that is, $B^2 = A$).

Even though the question explicitly mentions matrices, it is just as easy to discuss endomorphisms of the vector space $V = \mathbb{C}^n$.

By the spectral theorem, A is diagonalizable, so $V = \mathbb{C}^n$ is the sum of the eigenspaces V_λ of A. By hermitianness these eigenspaces are mutually orthogonal. By positive-definiteness A has *positive* real eigenvalues λ, which therefore have real square roots. Define B on each orthogonal summand V_λ to be the scalar $\sqrt{\lambda}$. Since these eigenspaces are mutually orthogonal, the operator B so defined really is hermitian, as we now verify. Let $v = \sum_\lambda v_\lambda$ and $w = \sum_\mu w_\mu$ be *orthogonal* decompositions of two vectors into eigenvectors v_λ with eigenvalues λ and w_μ with eigenvalues μ. Then, using the orthogonality of eigenvectors with distinct eigenvalues,

$$\langle Bv, w\rangle = \langle B\sum_\lambda v_\lambda, \sum_\mu w_\mu\rangle = \langle \sum_\lambda \lambda v_\lambda, \sum_\mu w_\mu\rangle = \sum_\lambda \lambda\langle v_\lambda, w_\lambda\rangle$$

$$= \sum_\lambda \langle v_\lambda, \lambda w_\lambda\rangle = \langle \sum_\mu v_\mu, \sum_\lambda \lambda w_\lambda\rangle = \langle v, Bw\rangle$$

Uniqueness is slightly subtler. Since we do not know *a priori* that two positive-definite square roots B and C of A *commute*, we *cannot* immediately say that $B^2 = C^2$ gives $(B+C)(B-C) = 0$, etc. If we *could* do that, then since B and C are both positive-definite, we could say

$$\langle (B+C)v, v\rangle = \langle Bv, v\rangle + \langle Cv, v\rangle > 0$$

so $B + C$ is positive-definite and, hence invertible. Thus, $B - C = 0$. But we cannot directly do this. We must be more circumspect.

Let B be a positive-definite square root of A. Then B commutes with A. Thus, B stabilizes each eigenspace of A. Since B is diagonalizable on V, it is diagonalizable on each eigenspace of A (from an earlier example). Thus, since all eigenvalues of B are *positive*, and $B^2 = \lambda$ on the λ-eigenspace V_λ of A, it must be that B is the scalar $\sqrt{\lambda}$ on V_λ. That is, B is uniquely determined. ///

[24.8] Given a square n-by-n complex matrix M, show that there are unitary matrices A and B such that AMB is *diagonal*.

We prove this for not-necessarily square M, with the unitary matrices of appropriate sizes.

This asserted expression

$$M = \text{unitary} \cdot \text{diagonal} \cdot \text{unitary}$$

is called a **Cartan decomposition** of M.

First, if M is *(square) invertible*, then $T = MM^*$ is self-adjoint and invertible. From an earlier example, the spectral theorem implies that there is a self-adjoint (necessarily invertible) square root S of T. Then

$$1 = S^{-1}TS^{-1} = (S^{-1}M)(^{-1}SM)^*$$

so $k_1 = S^{-1}M$ is unitary. Let k_2 be unitary such that $D = k_2 S k_2^*$ is diagonal, by the spectral theorem. Then

$$M = Sk_1 = (k_2 D k_2^*)k_1 = k_2 \cdot D \cdot (k_2^* k_1)$$

expresses M as

$$M = \text{unitary} \cdot \text{diagonal} \cdot \text{unitary}$$

as desired.

In the case of m-by-n (not necessarily invertible) M, we want to reduce to the invertible case by showing that there are m-by-m unitary A_1 and n-by-n unitary B_1 such that

$$A_1 M B_1 = \begin{pmatrix} M' & 0 \\ 0 & 0 \end{pmatrix}$$

where M' is *square* and invertible. That is, we can (in effect) do column and row reduction with *unitary* matrices.

Nearly half of the issue is showing that by left (or right) multiplication by a suitable unitary matrix A an arbitrary matrix M may be put in the form

$$AM = \begin{pmatrix} M_{11} & M_{12} \\ 0 & 0 \end{pmatrix}$$

with 0's below the r^{th} row, where the column space of M has dimension r. To this end, let $f_1, \ldots, f_r$ be an orthonormal basis for the *column space* of M, and extend it to an orthonormal basis $f_1, \ldots, f_m$ for the whole $\mathbb{C}^m$. Let $e_1, \ldots, e_m$ be the standard orthonormal basis for $\mathbb{C}^m$. Let A be the linear endomorphism of $\mathbb{C}^m$ defined by $Af_i = e_i$ for all indices i. We claim that this A is unitary, and has the desired effect on M. That it has the desired effect on M is by design, since any column of the original M will be mapped by A to the span of $e_1, \ldots, e_r$, so will have all 0's below the r^{th} row. A linear endomorphism is determined exactly by where it sends a basis, so all that needs to be checked is the unitariness, which will result from the orthonormality of the bases, as follows. For $v = \sum_i a_i f_i$ and $w = \sum_i b_i f_i$,

$$\langle Av, Aw \rangle = \langle \sum_i a_i \, Af_i, \sum_j b_j \, Af_j \rangle = \langle \sum_i a_i \, e_i, \sum_j b_j \, e_j \rangle = \sum_i a_i \overline{b_i}$$

by orthonormality. And, similarly,

$$\sum_i a_i \overline{b_i} = \langle \sum_i a_i \, f_i, \sum_j b_j \, f_j \rangle = \langle v, w \rangle$$

Thus, $\langle Av, Aw \rangle = \langle v, w \rangle$. To be completely scrupulous, we want to see that the latter condition implies that $A^*A = 1$. We have $\langle A^*Av, w \rangle = \langle v, w \rangle$ for all v and w. If $A^*A \neq 1$, then for some v we would have $A^*Av \neq v$, and for that v take $w = (A^*A - 1)v$, so

$$\langle (A^*A - 1)v, w \rangle = \langle (A^*A - 1)v, (A^*A - 1)v \rangle > 0$$

contradiction. That is, A is certainly unitary.

If we had had the foresight to prove that row rank is always equal to column rank, then we would know that a combination of the previous left multiplication by unitary and a corresponding right multiplication by unitary would leave us with

$$\begin{pmatrix} M' & 0 \\ 0 & 0 \end{pmatrix}$$

with M' *square* and invertible, as desired. ///

[24.9] Given a square n-by-n complex matrix M, show that there is a unitary matrix A such that AM is *upper triangular.*

Let $\{e_i\}$ be the standard basis for $\mathbb{C}^n$. To say that a matrix is upper triangular is to assert that (with left multiplication of column vectors) each of the maximal family of nested subspaces (called a **maximal flag**)

$$V_0 = 0 \subset V_1 = \mathbb{C}e_1 \subset \mathbb{C}e_1 + \mathbb{C}e_2 \subset \ldots \subset \mathbb{C}e_1 + \ldots + \mathbb{C}e_{n-1} \subset V_n = \mathbb{C}^n$$

is stabilized by the matrix. Of course

$$MV_0 \subset MV_1 \subset MV_2 \subset \ldots \subset MV_{n-1} \subset V_n$$

is another maximal flag. Let f_{i+1} be a unit-length vector in the orthogonal complement to MV_i inside MV_{i+1} Thus, these f_i are an orthonormal basis for V, and, in fact, $f_1, \ldots, f_t$ is an orthonormal basis for MV_t. Then let A be the unitary endomorphism such that $Af_i = e_i$. (In an earlier example and in class we checked that, indeed, a linear map which sends one orthonormal basis to another is unitary.) Then

$$AMV_i = V_i$$

so AM is upper-triangular. ///

[24.10] Let Z be an m-by-n complex matrix. Let Z^* be its conjugate-transpose. Show that

$$\det(1_m - ZZ^*) = \det(1_n - Z^*Z)$$

Write Z in the (rectangular) Cartan decomposition

$$Z = ADB$$

with A and B unitary and D is m-by-n of the form

$$D = \begin{pmatrix} d_1 & & & & & \\ & d_2 & & & & \\ & & \ddots & & & \\ & & & d_r & & \\ & & & & 0 & \\ & & & & & \ddots \end{pmatrix}$$

where the diagonal d_i are the only non-zero entries. We grant ourselves that $\det(xy) = \det(x) \cdot \det(y)$ for square matrices x, y of the same size. Then

$$\det(1_m - ZZ^*) = \det(1_m - ADBB^*D^*A^*) = \det(1_m - ADD^*A^*) = \det(A \cdot (1_m - DD^*) \cdot A^*)$$

$$= \det(AA^*) \cdot \det(1_m - DD^*) = \det(1_m - DD^*) = \prod_i (1 - d_i\overline{d_i})$$

Similarly,

$$\det(1_n - Z^*Z) = \det(1_n - B^*D^*A^*ADB) = \det(1_n - B^*D^*DB) = \det(B^* \cdot (1_n - D^*D) \cdot B)$$

$$= \det(B^*B) \cdot \det(1_n - D^*D) = \det(1_n - D^*D) = \prod_i (1 - d_i\overline{d_i})$$

which is the same as the first computation. ///

Exercises

24.[9.0.1] Let B be a bilinear form on a vector space V over a field k. Suppose that for $x, y \in V$ if $B(x, y) = 0$ then $B(y, x) = 0$. Show that B is either *symmetric* or *alternating*, that is, either $B(x, y) = B(y, x)$ for all $x, y \in V$ or $B(x, y) = -B(y, x)$ for all $x, y \in V$.

24.[9.0.2] Let R be a commutative ring of endomorphisms of a finite-dimensional vector space V over $\mathbb{C}$ with a hermitian inner product $\langle , \rangle$. Suppose that R is closed under taking adjoints with respect to $\langle , \rangle$. Suppose that the only R-stable subspaces of V are $\{0\}$ and V itself. Prove that V is one-dimensional.

24.[9.0.3] Let T be a self-adjoint operator on a complex vector space V with hermitian inner product $\bar{,}\rangle$. Let W be a T-stable subspace of V. Show that the restriction of T to W is self-adjoint.

24.[9.0.4] Let T be a diagonalizable k-linear endomorphism of a k-vectorspace V. Let W be a T-stable subspace of V. Show that T is diagonalizable on W.

24.[9.0.5] Let V be a finite-dimensional vector space over an algebraically closed field k. Let T be a k-linear endomorphism of V. Show that T can be written uniquely as $T = D + N$ where D is diagonalizable, N is nilpotent, and $DN = ND$.

24.[9.0.6] Let S, T be commuting k-linear endomorphisms of a finite-dimensional vector space V over an algebraically closed field k. Show that S, T have a common non-zero eigenvector.