Operators on Inner-Product Spaces

Self-Adjoint: $T^* = T$.

Normal: $TT^* = T^*T$.

Positive: Self-adjoint and $\langle Tv, v\rangle \geq 0$ for all v.

Isometry: $||Tv|| = ||v||$ for all v.

Let V be a finite-dimensional inner-product space. We wish to characterize those linear operators $T : V \to V$ that have an orthonormal basis of eigenvectors.

Complex Spectral Theorem: When V is a complex vector space, V has an orthonormal basis of eigenvectors with respect to a linear operator T if and only if T is normal.

Proof: If V has an orthonormal basis of eigenvectors with respect to a linear operator T then T has a diagonal matrix representation A, which implies $m(T^*) = A^*$ with respect to the same basis. Since the matrix representations of T and T^* are diagonal, they commute, hence T commutes with T^*.

Conversely, suppose $TT^* = T^*T$. Since V is a complex vector space, T has an upper-triangular matrix representation A with respect to an orthonormal basis (method: apply Gram-Schmidt to an upper-triangular basis). It will suffice to show that this matrix is diagonal. It will further suffice to show that $||R_i||^2 = ||C_i||^2$ for each i. Let $\{u_1, \ldots, u_n\}$ be the orthonormal basis. We must show $||T^*e_i||^2 = ||Te_i||^2$ for all i. However, for any $v \in V$,

$$||T^*v||^2 = \langle T^*v, T^*v\rangle = \langle TT^*v, v\rangle = \langle T^*Tv, v\rangle =$$

$$= \langle Tv, Tv\rangle = ||Tv||^2.$$

Real Spectral Theorem: When V is a real vector space, V has an orthonormal basis of eigenvectors with respect to a linear operator T if and only if T is self-adjoint.

Proof: If V has an orthonormal basis of eigenvectors with respect to a real operator T then T has a diagonal matrix representation A which satisfies $A^T = A$. This implies that T is self-adjoint.

Conversely, suppose that a real operator $T : V \to V$ is self-adjoint. We will define a complex normal operator $S : W \to W$, use the Complex Spectral

Theorem to find an orthonormal basis of eigenvectors for W, and use this to find an orthonormal basis of eigenvectors for V.

Let $W = V \times V$ with scalar multiplication defined by

$$(a + bi)(v, w) = (av - bw, aw + bv)$$

and inner product defined by

$$\langle (v_1, w_1), (v_2, w_2) \rangle = (\langle v_1, v_2 \rangle + \langle w_1, w_2 \rangle) + i(-\langle v_1, w_2 \rangle + \langle w_1, v_2 \rangle).$$

Let $S : W \to W$ be defined by

$$S(v, w) = (Tv, Tw).$$

S is self-adjoint, hence normal, hence $V \times V$ has an orthonormal basis of eigenvectors $\{(v_1, w_1), \ldots, (v_n, w_n)\}$. The equation $S(v, w) = (a + bi)(v, w)$ implies $T(v) = av - bw$ and $T(w) = aw + bv$. Making these substitutions into $\langle Tv, w \rangle = \langle v, Tw \rangle$ and rearranging yields $b(\langle v, v \rangle + \langle w, w \rangle) = 0$. This forces $b = 0$. Hence for each (v_i, w_i) there is a real number a_i such that $Tv_i = a_i v_i$ and $Tw_i = a_i w_i$. Since $\{(v_1, w_1), \ldots, (v_n, w_n)\}$ is a basis for $V \times V$ over $\mathbb{C}$, every $(x, 0)$ is in the span of $\{(v_1, w_1), \ldots, (v_n, w_n)\}$ over $\mathbb{C}$, which implies that each $x \in V$ is in the span of $\{v_1, w_1, \ldots, v_n, w_n\}$ over $\mathbb{R}$. Hence V has a basis which is a subset of $\{v_1, w_1, \ldots, v_n, w_n\}$, hence V has a basis of eigenvectors of T.

Now let B_a be the set of those basis eigenvectors corresponding to eigenvalue a. Applying Gram-Schmidt to these produces an orthonormal set of eigenvectors O_a corresponding to eigenvalue a. If u and v are eigenvectors corresponding to distinct eigenvalues of T then they are orthogonal to each other: suppose $Tu = au$ and $Tv = bv$ where $a \neq b$. Then

$$a\langle u, v \rangle = \langle au, v \rangle = \langle Tu, v \rangle =$$

$$\langle u, Tv \rangle = \langle u, bv \rangle = b\langle u, v \rangle.$$

This forces $\langle u, v \rangle = 0$. The union of the O_a bases forms an orthonormal eigenvector basis of V.

Remark: Taken together, the two spectral theorems yield an algorithm for finding an orthonormal basis of eigenvectors for an inner product space V given a linear operator $T : V \to V$ which is normal (V complex) or self-adjoint (V real): First, find all the eigenvalues of T. Second, compute each eigenspace. Since V has in principle an orthonormal basis of eigenvectors, V is the sum of its eigenspaces. Since eigenvectors corresponding to distinct eigenvalues are linearly independent, the sum is direct. Hence the union of the eigenspace bases forms a basis of eigenvectors for V. When V is complex and T is normal we know that applying Gram-Schmidt to a basis of eigenvectors produces an orthonormal basis of eigenvectors since the original basis produces an upper-triangular matrix representation. When V is real and T is self-adjoint we know that eigenvectors corresponding to distinct eigenvalues are orthogonal, so applying Gram-Schmidt to a eigenvector basis produces an orthonormal eigenvector basis. To summarize, in either case we find a basis for each eigenspace, then apply Gram-Schmidt to the union of these bases.

Characterization of Real Normal Linear Operators: Let $T : V \to V$ be a normal operator on a real inner product space V. We will show that T has a block-diagonal matrix representation with respect to an orthonormal basis, where each of the blocks has dimension ≤ 2 and each of the blocks of dimension 2 has the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. We can use an induction argument: for dimension 1 this is trivial. More generally, find an invariant subspace U of dimension ≤ 2, find an orthonormal basis for it, and expand to an orthonormal basis for V. The matrix representation of T restricted to U has the desired form, using the algebra steps which appear in the proof of the Complex Spectral Theorem. The same algebra steps can be used to show that the matrix representation of T with respect to this basis is $\begin{bmatrix} A_1 & 0 \\ 0 & B \end{bmatrix}$. So we have the foundation for an induction argument: we have decomposed V into $U \bigoplus U^\perp$ where T maps U into U and $U^\perp$ into $U^\perp$ and the matrix representation of T restricted to these invariant subspaces are A_1 and B and U has dimension ≤ 2. The induction hypothesis enables us to find an orthonormal basis for $U^\perp$ of the desired form, and adding these vectors to the orthonormal basis for U produces the desired orthonormal basis for V. Conversely, if T has such a matrix representation with respect to an orthonormal basis then it is normal. Note also that T is normal but not self-adjoint if and only if T has at least one 2×2 block of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ with $b \neq 0$. By replacing one basis vector with its additive inverse if necessary we can assume $b > 0$.

(Note: if T is represented by $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ with respect to the basis $\{u_1, u_2\}$ then it is represented by the matrix $\begin{bmatrix} a_{11} & (\beta/\alpha)a_{12} \\ (\alpha/\beta)a_{21} & a_{22} \end{bmatrix}$ with respect to the basis $\{\alpha u_1, \beta u_2\}$.)

Characterization of Positive Operators: An operator $T : V \to V$ on a complex or real finite-dimensional inner product space is said to be positive if it is self-adjoint and satisfies $\langle Tv, v\rangle \geq 0$ for each $v \in V$. We know that the self-adjoint operators are precisely those that have a diagonal matrix representation with respect to some orthonormal basis of eigenvectors, where the diagonal entries r_{ii} are real numbers. The equation

$$\langle T(\sum_i \alpha_i e_i), \sum_i \alpha_i e_i\rangle = \sum_i r_{ii}|\alpha_i|^2 \geq 0$$

holds for all choices of $(\alpha_1, \ldots, \alpha_n)$, which happens if and only if the diagonal entries are non-negative real numbers. So the positive operators are precisely those that have a diagonal matrix representation with respect to some orthonormal basis of eigenvectors, in which the diagonal entries (i.e. the eigenvalues) are all non-negative. Theorem 7.27 is a list of equivalent conditions for a positive operator T.

Given this characterization, each positive operator has a positive square root: just use the same orthonormal basis of eigenvectors, but let the matrix representation be that in which the the diagonal entries are replaced by their non-negative square roots. This square root is unique: let T be a positive operator and let S be a positive operator that satisfies $S^2 = T$. Computing S on the eigenvector basis yields $Sv = \sqrt{\lambda}v$ where $Tv = \lambda v$. This determines S. Notation: $S = \sqrt{T}$.

For any linear operator $T : V \to V$ on an inner product space V, the operator T^*T is always positive: $\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = ||Tv||^2 \geq 0$. Hence one can always construct $\sqrt{T^*T}$. The map $\sqrt{T^*T}$ has some interesting properties:

1. For all $u, v \in V$, $\langle\sqrt{T^*T}u, \sqrt{T^*T}v\rangle = \langle Tu, Tv\rangle$. Proof:

$$\langle\sqrt{T^*T}u, \sqrt{T^*T}v\rangle = \langle T^*Tu, v\rangle = \langle Tu, Tv\rangle.$$

2. $\text{null}(\sqrt{T^*T}) = \text{null}(T)$. Proof: First note that $\text{null}(T) \subseteq \text{null}(T^*T)$. On the other hand, if $u \in \text{null}(T^*T)$ then $T^*Tu = 0$, hence $\langle T^*Tu, u\rangle = 0$,

therefore $\langle Tu, Tu\rangle = 0$, therefore $||Tu||^2 = 0$, therefore $Tu = 0$, therefore $u \in \text{null}(T)$. Hence $\text{null}(T) = \text{null}(T^*T)$. Using the diagonal matrix representation of T^*T we can show that $\text{null}(T^*T) = \text{null}(\sqrt{T^*T})$. Hence $\text{null}(T) = \text{null}(\sqrt{T^*T})$.

3. $\text{rank}(\sqrt{T^*T}) = \text{rank}(T)$. Proof: This follows from #2 above and the Rank-Nullity Theorem.

Isometries: An operator $S : V \to V$ on a complex or real finite-dimensional inner product space is said to be an isometry if satisfies $||Sv|| = ||v||$ for each $v \in V$.

Theorem: The following are equivalent:

(a) $S : V \to V$ is an isometry.

(b) For all $u, v \in V$, $\langle Su, Sv\rangle = \langle u, v\rangle$.

(c) $S^*S = I$.

(d) S maps an orthonormal basis to an orthonormal basis.

Proof: (a) implies (b): Assume that S is an isometry. Expanding the equation

$$||S(\lambda u + v)||^2 = ||\lambda u + v||^2$$

yields

$$2\text{Re}(\lambda\langle Su, Sv\rangle) = 2\text{Re}(\lambda\langle u, v\rangle).$$

If V is real then (b) holds. If V is complex then setting $\lambda = 1$, then $\lambda = i$, we see that (b) holds.

(b) implies (c): Assume that S satisfies $\langle Su, Sv\rangle = \langle u, v\rangle$ for all $u, v \in V$. Then

$$\langle S^*Su, v\rangle = \langle u, v\rangle,$$

therefore

$$\langle (S^*S - I)u, v\rangle = 0.$$

Setting $v = (S^*S - I)u$ we obtain

$$||(S^*S - I)u|| = 0$$

for all $u \in V$, hence $S^*S = I$.

(c) implies (d): Assume that $S^*S = I$. Let $\{e_1, \ldots, e_n\}$ be an orthormal basis for V. Then

$$\langle Se_i, Se_j\rangle = \langle S^*Se_i, e_j\rangle = \langle e_i, e_j\rangle$$

for all i and j, hence $\{Se_1, \ldots, Se_n\}$ is orthonormal.

(d) implies (a): Assume that S maps the orthonormal basis $\{e_1, \ldots, e_n\}$ into the orthonormal basis $\{f_1, \ldots, f_n\}$. Then $\langle Se_i, Se_j\rangle = \langle e_i, e_j\rangle$ for all i and i, which implies $\langle Su, Sv\rangle = \langle u, v\rangle$ for all $u, v \in V$. This implies $||Sv|| = ||v||$ for all v.

Note that $S^*S = I$ implies that isometries are normal. We have already characterized normal operators above: complex normal operators have a diagonal matrix representation with respect to an orthonormal basis of eigenvectors, and real normal operators have a block-diagonal matrix representation in which the 2×2 blocks are of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. The equation $SS^* = I$ implies that the 1-dimensional blocks have norm 1 and the 2-dimensional blocks can be assumed to be in the form $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ for some $\theta \in (0, \pi)$ (where some of the othornormal basis vectors may need to be replaced by their additive inverses). Conversely, if a linear operator as a matrix representation of this form with respect to an orthonormal basis then it satisfes $S^*S = I$ hence it is an isometry.

Polar Decomposition and Singular Decomposition

Theorem 7.41 (Polar Decomposition of a Linear Operator): Let T be a linear operator on a finite-dimensional inner product space V. Then there is an isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

Proof: The operator $\sqrt{T^*T}$ maps V onto range($\sqrt{T^*T}$). The mapping $S_0 : \text{range}(\sqrt{T^*T}) \to \text{range}(T)$ defined by

$$S_0(\sqrt{T^*T}v) = Tv$$

is well-defined, because if

$$\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$$

then

$$\sqrt{T^*T}(v_1 - v_2) = 0,$$

therefore
$$\langle\sqrt{T^*T}(v_1 - v_2), \sqrt{T^*T}(v_1 - v_2)\rangle = 0,$$
therefore
$$\langle T(v_1 - v_2), T(v_1 - v_2) = 0,$$
therefore
$$Tv_1 = Tv_2.$$
Moreover,
$$\langle S_0(\sqrt{T^*T}u), S(\sqrt{T^*T}v)\rangle = \langle Tu, Tv\rangle = \langle\sqrt{T^*T}u, \sqrt{T^*T}v\rangle.$$
Hence S_0 is an inner-product preserving linear map. Since $\text{rank}(\sqrt{T^*T}) = \text{rank}(T)$, $\text{range}(\sqrt{T^*T})^\perp$ and $\text{range}(T)^\perp$ have the same dimension. Let $\{e_1, \ldots, e_k\}$ be an orthonormal basis for $\text{range}(\sqrt{T^*T})^\perp$ and let $\{f_1, \ldots, f_k\}$ be an orthonormal basis for $\text{range}(T)^\perp$. We will extend S_0 to a linear map $S : V \to V$ by declaring that $Se_i = f_i$ for each i. By construction, $\langle Su, Sv\rangle = \langle u, v\rangle$ for each $u, v \in V$, therefore S is an isometry. By construction, $T = S\sqrt{T^*T}$.

Note that if we choose an orthonormal basis of eigenvectors for $\sqrt{T^*T}$ of the form $\{e_1, \ldots, e_n\}$ then we have $T(e_i) = S\sqrt{T^*T}e_i = Ss_ie_i = s_iSe_i$. Setting $f_i = Se_i$ for each i, the matrix representation of T with respect to the pair of orthonormal bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_n\}$ is $\text{diag}(s_1, \ldots, s_n)$. This choice of bases is called the Singular Value Decomposition of T.

Miscellaneous Results:

Proposition 7.1: Every eigenvalue of a self-adjoint operator is real.

Proof: Let $T : V \to V$ be self-adjoint and V a complex vector space. Then T is normal and has a diagonal matrix representation A. Since $A = A^*$, its diagonal entries are real, hence the eigenvalues of T are real.

Proposition 7.3: Let $T : V \to V$ be a linear operator on a complex inner-product space. Then T is self-adjoint if and only if $\langle Tv, v\rangle \in \mathbb{R}$ for all $v \in V$.

Proof: Assume T is self-adjoint. Then it is normal, hence has a diagonal matrix representation with real diagonal entries r_{ii} with respect to an orthonormal basis of eigenvectors. This implies
$$\langle T(\sum_i \alpha_i e_i), \sum_i \alpha_i e_i\rangle = \sum_i r_{ii}|\alpha_i|^2 \in \mathbb{R}.$$

Conversely, if $\langle Tv, v\rangle \in \mathbb{R}$ for all $v \in V$ then

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

hence T is self-adjoint.

Proposition 7.4: If T is a self-adjoint operator on V such that $\langle Tv, v\rangle = 0$ for all $v \in V$ then $T = 0$.

Proof: T has a diagonal matrix representation with respect to an orthonormal basis of eigenvectors. Then

$$\langle T(\sum_i \alpha_i e_i), \sum_i \alpha_i e_i\rangle = \sum_i r_{ii}|\alpha_i|^2 = 0$$

holds for all choices of $(\alpha_1, \ldots, \alpha_n)$. This implies that all the eigenvalues r_{ii} are 0, which implies that T is represented by the 0 matrix, which says $T = 0$.

Proposition 7.6: A linear operator $T : V \to V$ on an inner-product space is normal if and only if $||Tv|| = ||T^*v||$ for all v.

Proof: T is normal iff $T^*T - TT^* = 0$ iff $\langle (T^*T - TT^*)v, v\rangle = 0$ iff $\langle T^*Tv, v\rangle = \langle TT^*v, v\rangle$ iff $\langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle$ iff $||Tv|| = ||T^*v||$ for all v.

Corollary 7.7: Let $T : V \to V$ be normal. Then v is an eigenvector of T with respect to eigenvalue λ if and only if v is an eigenvector of T^* with respect to eigenvalue $\overline{\lambda}$.

Proof: Let v be an eigenvector of T. Then $||(T - \lambda I)v|| = 0$. Since T is normal, $T - \lambda I$ is normal with $(T - \lambda I)^* = T^* - \overline{\lambda} I$. Hence $||(T^* - \overline{\lambda})v|| = 0$. This makes v an eigenvector of T^* with eigenvalue $\overline{\lambda}$. The converse holds as well.

Proposition 7.8: Let $T : V \to V$ be normal. Eigenvectors corresponding to distinct eigenvalues of T are orthogonal.

Proof: Suppose $Tu = \alpha u$ and $Tv = \beta v$ where $\alpha \neq \beta$. Then

$$\begin{aligned}\alpha\langle u, v\rangle &= \langle Tu, v\rangle = \langle u, T^*v\rangle \\ &= \langle u, \overline{\beta}v\rangle = \beta\langle u, v\rangle.\end{aligned}$$

Hence $\langle u, v\rangle = 0$.