6C Orthogonal Complements and Minimization Problems

- V and W continue to denote inner product spaces over $\mathbf{F}$.

Orthogonal Complements

- Definition 6.46: The **orthogonal complement** of any subset $U \subseteq V$ is the subset $U^\perp = \{v \in V \mid u \perp v \text{ for all } u \in U\}$.
  - We are mainly interested in the case in which U is a subspace.
- Examples 6.47:
  - The orthogonal complement of a line/plane through the origin in $\mathbf{R}^3$ is a plane/line through the origin.
  - In $\mathbf{F}^n$, the orthogonal complement of $\operatorname{span}(e_1, \ldots, e_k)$ is $\operatorname{span}(e_{k+1}, \ldots, e_n)$.
  - This holds more generally whenever $e_1, \ldots, e_n$ is an orthonormal basis of V.
- Proposition 6.48:
  (a) $U^\perp$ is a subspace
  (b) $\{0\}^\perp = V$
  (c) $V^\perp = \{0\}$
  (d) $U \cap U^\perp \subseteq \{0\}$
  (e) if $U_1 \subseteq U_2$, then $U_2^\perp \subseteq U_1^\perp$.
- Proposition 6.49: if $U \subseteq V$ is a finite-dimensional subspace, then $V = U \oplus U^\perp$.
- Definition 6.55: if $U \subseteq V$ is a finite-dimensional subspace, then the **orthogonal projection** onto U is the operator $P_U \in \mathcal{L}(V)$ such that $P_U(u + w) = u$ for $u \in U$ and $w \in U^\perp$.
  - Explicitly, if $u_1, \ldots, u_k$ is an orthogonal basis of U, then $P_U(w) = \sum_{i=1}^k \frac{\langle w, u_i \rangle}{\langle u_i, u_i \rangle} u_i$.
- Corollary 6.51: if V is finite-dimensional and $U \subseteq V$ is a subspace, then $\dim U^\perp = \dim V - \dim U$.
- Corollary 6.52: if $U \subseteq V$ is a finite-dimensional subspace, then $(U^\perp)^\perp = U$.
- Corollary 6.54: if $U \subseteq V$ is a finite-dimensional subspace, then $U^\perp = \{0\} \iff U = V$.
- Proposition 6.57: if U is a finite-dimensional subspace of V, then (i) P_U is a projection operator onto U (see below), and (ii) $\|P_U v\| \leq \|v\|$ for all $v \in V$.

Aside on projection operators

- This is something that we should have covered when we discussed direct sums; during this aside, let V be an arbitrary vector space (rather than an inner product space).
- Definition: given a subspace $U \subseteq V$ of a vector space V, a **projection operator** onto U is an operator $P \in \mathcal{L}(V)$ with $\operatorname{range}(P) = U$ and $P|_U = I$.
- Proposition: $P \in \mathcal{L}(V)$ is a projection operator if and only if $P^2 = P$ (an operator with this property is also called **idempotent**), in which case it is a projection operator onto $U = \operatorname{range}(P)$.
  - Proof: if P is a projection onto $U \subseteq V$, then $P(v) \in U$ for any $v \in V$, hence $P(P(v)) = P(v)$.
  - Conversely, if $P^2 = P$, we may set $U = \operatorname{range} P$; then any $u \in U$ is of the form $u = Pv$, and hence $Pu = PPv = Pv = u$.
- Definition: given a subspace $U \subseteq V$, a **complement** of U is a subspace $W \subseteq V$ with $U \oplus W = V$.
- Theorem: let $U \subseteq V$ be a subspace. Then
  (a) if P is a projection onto U, then null P is a complement of U;
  (b) conversely, if W is any complement of U, there is a unique projection P onto U with null $P = W$.
  - Proof of (a): if $u \in U \cap \operatorname{null} P$, then $u = Pu = 0$, where the first equality follows from $P|_U = I$; and given any $v \in V$, we have $P(v - Pv) = Pv - PPv = Pv - Pv = 0$, hence $v - Pv \in \operatorname{null} P$, and so $v = Pv + (v - Pv) \in U + \operatorname{null} P$.
  - Proof of (b): we define a projection onto U as follows: given $v \in V$, we can write v uniquely as $v = u + w$ with $u \in U$ and $w \in W$, and we set $Pv = u$; it's easy to check that this is a projection on to U with nullspace W. To prove uniqueness, given any projection P onto U with nullspace W, if $v = u + w$ as above, then we must have $Pv = Pu + Pw = u + 0 = u$.
- Proposition: if P is a projection onto U, then $I - P$ is a projection onto the complement null P of U.
  - Proof: we have $(I-P)^2 = I^2 - 2P + P^2 = I - 2P + P = I - P$, hence $I - P$ is a projection operator; it remains to see that $\operatorname{range}(I - P) = \operatorname{null} P$.
  - For any $v \in V$, we have $P((I-P)v) = P(v - Pv) = Pv - PPv = 0$, so $\operatorname{range}(I-P) \subseteq \operatorname{null} P$.
  - And for any $v \in \operatorname{null} P$, we have $v = v - Pv = (I-P)v \in \operatorname{range}(I-P)$, so $\operatorname{null} P \subseteq \operatorname{range}(I-P)$.

Orthogonal complements and annihilators

- We return to our convention that V and W denote inner product spaces over $\mathbf{F}$.
- Proposition: if V is finite-dimensional, then for any subspace $U \subseteq V$, the isomorphism $\overline{V} \to V'$ defined in the proof of (6.42) satisfies $F(U^\perp) = U^0$.
  - This follows directly from the definition: $v \in U^\perp$ if and only if $\langle u, v \rangle = 0$ for all $u \in U$ if and only if $F(v)|_U = 0$ if and only if $F(v) \in U^0$.

Minimization Problems

- Proposition 6.61: if U is a finite-dimensional subspace of V, then for any $v \in V$, $P_U v$ is the closest point to v in U, in the sense that $\|v - P_U v\| < \|v - u\|$ for any other $u \in U$.
- Example 6.63: we can use this to approximate the function $\sin\colon [-\pi, \pi]$ by a polynomial by letting $V = \mathcal{C}^0([-\pi, \pi])$ with inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} fg$, and letting $U = \operatorname{span}(1, x, \ldots, x^n)$.

7A Self-Adjoint and Normal Operators

- V and W continue to denote inner product spaces over $\mathbf{F}$.
- We will now introduce self-adjoint operators—which are an analogue for operators of symmetric matrices—and normal operators, which are a generalization of self-adjoint operators.
- The most important fact about these, which we will prove in the next section, is the (finite-dimensional) **spectral theorem** (Theorems 7.29 and 7.31), which says that

  (a) every self-adjoint operator T on a finite-dimensional vector space V is diagonalizable with respect to an orthonormal basis,

  (b) if $\mathbf{F} = \mathbf{C}$, then more generally, every **normal** operator T is diagonalizable with respect to an orthonormal basis,

  (c) conversely, if an operator is diagonalizable with respect to an orthonormal basis then, if $\mathbf{F} = \mathbf{R}$, it is self-adjoint, and if $\mathbf{F} = \mathbf{C}$, it is normal.

Adjoints

- Proposition/Definition 7.1 and 7.4: for any $T \in \mathcal{L}(V, W)$, there exists at most one linear map $T^*\colon W \to V$ satisfying $\langle Tv, w\rangle = \langle v, T^*w\rangle$ for all $v \in V$ and $w \in W$; if it exists, T^* is called the **adjoint** of T. Moreover, if $\dim V, \dim W < \infty$, then the adjoint T^* always exists.
- Remark: supposing $\dim V, \dim W < \infty$, let $\varphi_V\colon \overline{V} \xrightarrow{\sim} V'$ and $\varphi_W\colon \overline{W} \xrightarrow{\sim} W'$ be the isomorphisms given by $\varphi_V(v) = \langle -, v\rangle$ and $\varphi_W(w) = \langle -, w\rangle$. Then for any $T\colon \mathcal{L}(V, W)$ and $w \in \overline{W} = W$, we have $\varphi_V(T^*w) = T'(\varphi_W w)$.

$$\begin{array}{ccc} \overline{W} & \xrightarrow{T^*} & \overline{V} \\ {\scriptstyle\varphi_W}\downarrow\wr & & \wr\downarrow{\scriptstyle\varphi_V} \\ W' & \xrightarrow{T'} & V' \end{array}$$

  - Thus, roughly speaking, the adjoint of a linear map between inner product spaces is "isomorphic" to the dual map; in particular, all the properties of the dual map carry over to give properties of the adjoint map.
- Example: for a linear map $T\colon \mathbf{R}^{m,1} \to \mathbf{R}^{n,1}$ (with the Euclidean inner product) given by multiplying by a matrix A, the adjoint $T^*\colon \mathbf{R}^{n,1} \to \mathbf{R}^{m,1}$ is given by multiplying by A^{t}.
  - Indeed, one way of writing the Euclidean inner product on $\mathbf{R}^{n,1}$ is as the matrix product $\langle u, v\rangle = u^{\mathrm{t}} \cdot v \in \mathbf{R}^{1,1}$.

– Hence $\langle A \cdot u, v\rangle = (A \cdot u)^{\mathrm{t}} \cdot v = u^{\mathrm{t}} \cdot A^{\mathrm{t}} \cdot v = u^T \cdot (A^{\mathrm{t}} \cdot v) = \langle u, A^{\mathrm{t}} \cdot v\rangle$.

- Proposition 7.5: Suppose V and W are finite-dimensional. Then
  (a,b) the map $T \mapsto T^*$ is an anti-linear map $\mathcal{L}(V, W) \to \mathcal{L}(W, V)$
  (c) $(T^*)^* = T$ for any $T \in \mathcal{L}(V, W)$
  (d) $(ST)^* = T^* S^*$ for $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$ (where U is a finite-dimensional inner-product space over $\mathbf{F}$)
  (e) $I_V^* = I_V$
  (f) if $T \in \mathcal{L}(V, W)$ is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

- Proposition 7.6: if V and W are finite-dimensional, then for any $T \in \mathcal{L}(V, W)$:
  (a) null $T^* = (\operatorname{range} T)^\perp$
  (b) $\operatorname{range} T^* = (\text{null } T)^\perp$
  (c) null $T = (\text{null } T^*)^\perp$
  (d) $\operatorname{range} T = (\text{null } T^*)^\perp$

- Definition 7.7: the **conjugate transpose** of a matrix $A \in \mathbf{F}^{m,n}$ is the matrix A^* defined by $(A^*)_{j,k} = \overline{A_{k,j}}$.

- Proposition 7.9: given $T \in \mathcal{L}(V, W)$ and orthonormal bases for V and W, we have $\mathcal{M}(T^*) = \overline{\mathcal{M}(T)}$.

Self-Adjoint Operators

- Definition 7.10: $T \in \mathcal{L}(V)$ is **self-adjoint** if $T = T^*$, i.e., if $\langle Tu, v\rangle = \langle u, Tv\rangle$ for all $u, v \in V$.
  - Remark: given an orthonormal basis V, T is self-adjoint if and only if $\mathcal{M}(T)$ is **conjugate-symmetric** (also known as being **Hermitian**), meaning that $\mathcal{M}(T)^* = \mathcal{M}(T)$.

- Proposition 7.12: Every eigenvalue of a self-adjoint operator is real.
  - Of course, this is only an interesting result when $\mathbf{F} = \mathbf{C}$.

- Proposition 7.13: if $\mathbf{F} = \mathbf{C}$, then for any $T \in \mathcal{L}(V)$, we have $Tv \perp v$ for all $v \in V$ if and only if $T = 0$.
  - This is not true when $\mathbf{F} = \mathbf{R}$, for instance when $T(x, y) = (-y, x)$, a rotation by 90°.

- Proposition 7.14: if $F = \mathbf{C}$, then $T \in \mathcal{L}(V)$ is self-adjoint if and only if $\langle Tv, v\rangle \in \mathbf{R}$ for all $v \in V$.
  - This is obviously false when $\mathbf{F} = \mathbf{R}$, since there are real operators that are not self adjoint.

- Proposition 7.16: if $T \in \mathcal{L}(V)$ is self-adjoint, then $Tv \perp v$ for all $v \in V$ if and only if $T = 0$.
  - When $\mathbf{F} = \mathbf{C}$, this is implied by the stronger result (7.13), which does not have the self-adjoint hypothesis.

Normal operators

- Definition 7.18: $T \in \mathcal{L}(V)$ is **normal** if T and T^* commute.
  - Note that every self-adjoint operator is normal.
  - The main significance of normal operators is the **spectral theorem** in section 7B: if $F = \mathbf{C}$ and $\dim V < \infty$, then T is normal if and only if there is an orthonormal basis $\mathcal{B}$ such that $\mathcal{M}(T, \mathcal{B})$ is diagonal.
- Example 7.19: the operator on $\mathbf{F}^2$ with matrix $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$ is normal but not self-adjoint.
- Proposition 7.20: $T \in \mathcal{L}(V)$ is normal iff $\|Tv\| = \|T^*v\|$ for all $v \in V$.
- Proposition 7.21: if $T \in \mathcal{L}(V)$ is normal, then (a) null $T =$ null T^*
  (b) $\operatorname{range} T = \operatorname{range} T^*$ assuming $\dim V < \infty$
  (c) $V = \text{null } T \oplus \operatorname{range} T$ assuming $\dim V < \infty$
  (d) $T - \lambda I$ is normal for all $\lambda \in \mathbf{F}$
  (e) $Tv = \lambda v$ iff $T^*v = \bar{\lambda} v$ for $v \in V$ and $\lambda \in F$.
- Proposition 7.22: if $T \in \mathcal{L}(V)$ is normal, then eigenvectors of T with distinct eigenvalues are orthogonal.
  - This has a simpler proof in the special case that T is self-adjoint.
- Proposition 7.23: when $F = \mathbf{C}$, then $T \in \mathcal{L}(V)$ is normal if and only if there are self-adjoint commuting $A, B \in \mathcal{L}(V)$ with $T = A + iB$.
  - A version of the statement that makes sense over $\mathbf{R}$ is: T is normal if and only if there are commuting $A, B \in \mathcal{L}(V)$ with $T = A + B$ and $A = A^*$ and $B = -B^*$.

7B Spectral Theorem

- V and W continue to denote inner product spaces over $\mathbf{F}$.

Real Spectral Theorem

- Lemma 7.26: if $\dim V < \infty$ and $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbf{R}$ satisfy $b^2 - 4c < 0$, then $T^2 + bT + cI \in \mathcal{L}(V)$ is invertible.
- Lemma 7.27: if $\dim V < \infty$ and $T \in \mathcal{L}(V)$ is self-adjoint, then the minimal polynomial m_T is a product of real linear factors.
- Theorem 7.29 (real finite-dimensional spectral theorem): if $\mathbf{F} = \mathbf{R}$ and $\dim V < \infty$, then any $T \in \mathcal{L}(V)$ is self-adjoint if and only if it is diagonalizable with respect to some orthonormal basis.
  - Remark: an equivalent condition to being diagonalizable with respect to an orthonormal basis is that the eigenspaces of T are mutually orthogonal and their sum is V.

Complex Spectral Theorem

- Theorem 7.31 (complex finite-dimensional spectral theorem): if $\mathbf{F} = \mathbf{C}$ and $\dim V < \infty$, then any $T \in \mathcal{L}(V)$ is normal if and only if it is diagonalizable with respect to some orthonormal basis.