MAT 247S - Orthogonal projections

Let V be an inner product space. Recall that if S is a nonempty subset of V, then we define the orthogonal complement of S in V to be the set $S^\perp = \{ x \in V \mid \langle x, y \rangle = 0 \text{ for all } y \in S \}$. It is an easy exercise to prove that $S^\perp$ is a subspace of V.

Lemma. *Let S be a subset of V that contains $\mathbf{0}$. Then $S \cap S^\perp = \{\mathbf{0}\}$.*

Proof. Let $x \in S \cap S^\perp$. Then $\langle x, x \rangle = 0$. According to the fourth property of inner products, we must have $x = \mathbf{0}$.

Lemma. *Let S be a subset of V. Let $W = \operatorname{span}(S)$. Then $S^\perp = W^\perp$.*

Proof. Since $S \subset W$, it follows from the definitions of $S^\perp$ and $W^\perp$ that $W^\perp \subset S^\perp$.

Let $y \in W$. Then there exist vectors $y_1, \dots, y_n \in S$ and scalars $c_1, \dots, c_n$ such that $y = c_1y_1 + \cdots + c_ny_n$. Then, using the properties of inner products, we have, for $x \in S^\perp$,

$$\langle y, x \rangle = \sum_{j=1}^{n} \langle c_jy_j, x \rangle = \sum_{j=1}^{n} c_j \langle y_j, x \rangle = 0,$$

since $y_j \in S$ and $x \in S^\perp$. Therefore we have $\langle x, y \rangle = \overline{\langle y, x \rangle} = \bar{0} = 0$ for all $x \in S^\perp$ and all $y \in W$. This tells us that $S^\perp \subset W^\perp$. Since we already had the reverse inclusion, the lemma follows.

Corollary. *If W is a finite-dimensional subspace of an inner product space V and $\{ y_1, \dots, y_d \}$ is a basis of W, then $W^\perp = \{ x \in V \mid \langle x, y_j \rangle = 0, \text{ for } 1 \le j \le d \}$.*

If W is a finite-dimensional subspace of an inner product space V, the linear operator $T \in \mathcal{L}(V)$ described in the next theorem will be called the *orthogonal projection of V on W* (see the first paragraph on page 399 of the text, and also Theorem 6.6 on page 350).

Theorem. *Let W be a finite-dimensional subspace of an inner product space V.*

(1) There exists a unique $T \in \mathcal{L}(V)$ such that $T(x) = x$ for all $x \in W$ and $W^\perp = N(T)$.

(2) Suppose that V is finite-dimensional. Then $\dim W + \dim W^\perp = \dim V$.

(3) Let $\{ y_1, \dots, y_d \}$ be an orthonormal basis for W ($d = \dim W$). Let T be as in part (1). Then $T(x) = \sum_{j=1}^{d} \langle x, y_j \rangle y_j$, $x \in V$.

Proof. Let $\{ y_1, \dots, y_d \}$ be an orthonormal basis of W. (Note that such a basis must exist, according to Theorem 6.5.) Define $T(x) = \sum_{j=1}^{d} \langle x, y_j \rangle y_j$. First, we show that T is linear: For $x, z \in V$ and $c \in F$,

$$T(x+z) = \sum_{j=1}^{d} \langle x+z, y_j \rangle y_j = \sum_{j=1}^{d} (\langle x, y_j \rangle + \langle z, y_j \rangle) y_j = \sum_{j=1}^{d} (\langle x, y_j \rangle y_j + \langle z, y_j \rangle y_j) = T(x) + T(z)$$

$$T(cx) = \sum_{j=1}^{d} \langle cx, y_j \rangle y_j = \sum_{j=1}^{d} c \langle x, y_j \rangle y_j = c \left(\sum_{j=1}^{d} \langle x, y_j \rangle y_j \right) = cT(x).$$

Now suppose that $x \in W$. Then, according to Theorem 6.5, we must have $x = \sum_{j=1}^{d} \langle x, y_j \rangle y_j$. This is the same as $x = T(x)$. Therefore $T(x) = x$ for all $x \in W$.

Next, let $x \in W^\perp$. Then $\langle x, y_j \rangle = 0$ for $1 \le j \le d$, and so $T(x) = \sum_{j=1}^{d} 0 \cdot y_j = \{\mathbf{0}\}$. Therefore $W^\perp \subset N(T)$.

Now take $x \in N(T)$. By definition, we have $T(x) = \mathbf{0}$. It follows from linear independence of the vectors $\{ y_1, \dots, y_d \}$ that $\sum_{j=1}^{d} \langle x, y_j \rangle y_j = \mathbf{0}$ implies $\langle x, y_j \rangle = 0$ for $1 \le j \le d$. According to the corollary above, since $\{ y_1, \dots, y_d \}$ is a basis of W, we have $x \in W^{\perp}$. This shows $N(T) \subset W^{\perp}$.

To prove uniqueness of T, suppose $U \in \mathcal{L}(V)$ has the properties $U(x) = x$ for all $x \in W$ and $N(U) = W^{\perp}$. We need to show that $U = T$. Using linearity of U, we have

$$U(x) = U(x - T(x) + T(x)) = U(x - T(x)) + U(T(x)), \qquad x \in V.$$

Note that

$$\langle x - T(x), y_k \rangle = \langle x, y_k \rangle - \sum_{j=1}^{d} \langle x, y_j \rangle \langle y_j, y_k \rangle = \langle x, y_k \rangle - \langle x, y_k \rangle = 0,$$

$x - T(x) \in W^{\perp}$. Therefore $x - T(x) \in N(U)$: $U(x - T(x)) = \mathbf{0}$. Also, by definition of T, we have $T(x) \in \operatorname{span}\{ y_1, \dots, y_d \} = W$. Thus $U(T(x)) = T(x)$. It follows that $U(x) = \mathbf{0} + T(x) = T(x)$.

Now assume that V is finite-dimensional. Note that the property $T(x) = x$ for all $x \in W$ shows that $W \subset R(T)$. But we already know that $R(T) \subset W$ because $T(x) \in W$ for all $x \in V$. Therefore $R(T) = W$. Now use $N(T) = W^{\perp}$, together with the dimension theorem to see that part (3) holds.

For the last part, note that the uniqueness property of T proved in part (1) tells us that part (4) holds regardless of which orthonormal basis of W is used.

Remark. If W is a finite-dimensional subspace of V and T is the orthogonal projection of V on W, the proof of the theorem shows that $x - T(x) \in W^{\perp}$, and since $T(x) \in W$ by definition, we can write $x = T(x) + (x - T(x))$ and get a decomposition of x as the sum of a vector in W and a vector in $W^{\perp}$, as in Theorem 6.6 of the text.

Lemma. *If W and T are as in the theorem, then T is an orthogonal projection in the sense of the definition on page 398 of the text.*

Proof. The way T is defined, we have that $N(T) = W^{\perp}$ and $R(T) = W$ (see the proof above). This tells us that $N(T) = R(T)^{\perp}$. Since $W = (W^{\perp})^{\perp}$ (check this as an exercise), we also have $R(T) = (W^{\perp})^{\perp} = N(T)^{\perp}$, and the definition is satisfied.