Solutions to Homework Set #3

1. **Orthogonal complement of a subspace.** Suppose that $V$ is a subspace of $\mathbb{F}^n$. Let

$$V^\perp = \{x \in \mathbb{F}^n : x'y = 0, \forall y \in V\}$$

be the set of vectors orthogonal to every element in $V$.

(a) Verify that $V^\perp$ is a subspace of $\mathbb{F}^n$.

(b) Suppose that $V = \text{span}(v_1, v_2, \ldots, v_k)$ for some $v_1, v_2, \ldots, v_k \in \mathbb{F}^n$. Express $V$ and $V^\perp$ as subspaces induced by the matrix $A = [v_1 \ v_2 \ \cdots \ v_k] \in \mathbb{F}^{n \times k}$ and its transpose $A'$.

(c) Show that $(V^\perp)^\perp = V$.

(d) Show that $\dim(V) + \dim(V^\perp) = n$.

(e) Show that $V \subseteq W$ for another subspace $W$ implies $W^\perp \subseteq V^\perp$.

(f) Suppose that $\mathbb{F} = \mathbb{R}$. Show that every $x \in \mathbb{F}^n$ can be expressed uniquely as $x = v + v^\perp$, where $v \in V$ and $v^\perp \in V^\perp$. (Hint: Let $v$ be the projection of $x$ on $V$.)

**Solution:**

(a) We have, for all $y \in V$, $0'y = 0$ and therefore, $0 \in V^\perp$. Now, let $a, b \in \mathbb{F}$ and $u_1, u_2 \in V^\perp$. Then, we have, for all $y \in V$,

$$(au_1 + bu_2)'y = au_1'y + bu_2'y = 0,$$

since $u_1'y = u_2'y = 0$. Thus, $V^\perp$ is a subspace of $\mathbb{F}^n$.

(b) We have

$$V = \{x_1v_1 + \cdots + x_kv_k : x_1, x_2, \ldots, x_k \in \mathbb{F}\}$$

$$= \left\{ A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} : x_1, x_2, \ldots, x_k \in \mathbb{F} \right\}$$

$$= \{Ax : x \in \mathbb{F}^k\}$$

$$= \mathcal{R}(A),$$

and

$$V^\perp = \{y \in \mathbb{F}^n : y'(x_1v_1 + x_2v_2 + \cdots + x_kv_k) = 0, \forall x_1, \ldots, x_k \in \mathbb{F}\}$$

$$= \{y \in \mathbb{F}^n : (x_1v_1' + x_2v_2' + \cdots + x_kv_k')y = 0, \forall x_1, \ldots, x_k \in \mathbb{F}\}$$

$$= \{y \in \mathbb{F}^n : x'A'y = 0, \forall x \in \mathbb{F}^k\}$$

$$= \{y \in \mathbb{F}^n : A'y = 0\}$$

$$= \mathcal{N}(A').$$
(c) We will first show that $V \subseteq (V^\perp)^\perp$. Let $x \in V$. Then, by the definition of $V^\perp$, $x'y = 0$ for all $y \in V^\perp$ and therefore, $x \in (V^\perp)^\perp$. Thus, $V \subseteq (V^\perp)^\perp$. Now, by part (d), $(V^\perp)^\perp$ has dimension $\dim(V)$ and by what we just proved, includes $V$. This implies that $(V^\perp)^\perp = V$.

(d) From part (b), we have

\[
\dim(V) = \dim(\mathcal{R}(A)) = \text{rank}(A) = \text{rank}(A').
\]

Also,

\[
\dim(V^\perp) = \dim(V'(A')) = \text{nullity}(A').
\]

Therefore, since $A' \in \mathbb{F}^{k \times n}$, we have, by the rank–nullity theorem, $\dim(V) + \dim(V^\perp) = n$.

(e) Let $V \subseteq W$. We have

\[
x \in W^\perp \implies x'w = 0 \text{ for all } w \in W
\]

\[
\implies (a) x'w = 0 \text{ for all } w \in V
\]

\[
\implies x \in V^\perp.
\]

Here, implication $(a)$ follows since every $w \in V$ is also included in $W$, by assumption.

(f) Consider $v$, the projection of $x$ on $V$. Then, by the property of a projection, $v^\perp := x - v$ is orthogonal to $V$ and hence to all vectors in $V$. Therefore, $v^\perp \in V^\perp$. Now, suppose $x = v + v^\perp = \tilde{v} + \tilde{v}^\perp$, where $v, \tilde{v} \in V$ and $v^\perp, \tilde{v}^\perp \in V^\perp$. Then, we have $x = v^\perp = \tilde{v}^\perp = v^\perp$. But $v - \tilde{v} \in V$ and $\tilde{v}^\perp - v^\perp \in V^\perp$, therefore $v - \tilde{v} \in V \cap V^\perp = \{0\}$, implying that $v = \tilde{v}$, demonstrating the uniqueness of the representation.

Note that $V \cap V^\perp = \{0\}$ may not hold for subspaces over all fields $\mathbb{F}$. Consider, for example, the subspace $V$ of $\mathbb{F}_2^4$ spanned by $[1 \ 1 \ 1 \ 1]'$. Then,

\[
\mathcal{V} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\},
\]

and as can be verified easily,

\[
\mathcal{V}^\perp = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}.
\]

Now consider the vector $[1 \ 1 \ 0 \ 0]'$. We have

\[
\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix},
\]

demonstrating the non-uniqueness of the representation.
2. **Halfspace.** Suppose that \( a, b \in \mathbb{R}^n \) are two given points. Show that the set of points in \( \mathbb{R}^n \) that are closer to \( a \) than \( b \) is a halfspace, i.e.,

\[
\{ x : \| x - a \| \leq \| x - b \| \} = \{ x : c'x \leq d \}
\]

for appropriate \( c \in \mathbb{R}^n \) and \( d \in \mathbb{R} \).

(a) Find \( c \) and \( d \) explicitly in terms of \( a \) and \( b \).

(b) Draw a picture showing \( a, b, c, \) and the halfspace.

**Solution:**

(a) We have

\[
\{ x : \| x - a \| \leq \| x - b \| \} = \{ x : \| x - a \| ^2 \leq \| x - b \| ^2 \}
\]

\[
= \{ x : (x - a)'(x - a) \leq (x - b)'(x - b) \}
\]

\[
= \{ x : x'x - a'x - x'a + a'a \leq x'x - b'x - x'b + b'b \}
\]

\[
= \{ x : -2a'x + a'a \leq -2b'x + b'b \}
\]

\[
= \{ x : 2(b - a)'x \leq b'b - a'a \}
\]

\[
= \{ x : c'x \leq d \},
\]

where \( c := b - a \) and \( d := (|b|^2 - |a|^2)/2 \).

(b) 

3. **Inner product of polynomials.** Let \( P_3 \) be the vector space of all polynomials of degree \( \leq 3 \) with real coefficients, that is,

\[
P_3 = \{ \alpha_0 + \alpha_1x + \alpha_2x^2 + \alpha_3x^3 : \alpha_0, \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R} \}.
\]

Let \( K : P_3 \times P_3 \rightarrow \mathbb{R} \) be defined as

\[
K(p, q) = \int_{-1}^{1} p(x)q(x)dx.
\]
(a) Show that $K(\cdot, \cdot)$ represents an inner product for $\mathcal{P}_3$.

(b) Find an orthogonal basis for $\mathcal{P}_3$ using Gram–Schmidt orthogonalization.

Solution:

(a) We will show that the three properties that an inner product is required to satisfy hold for $K(\cdot, \cdot)$.

- **Linearity in the first argument.**

  $$K(\alpha p_1 + \beta p_2, q) = \int_{-1}^{1} (\alpha p_1(x) + \beta p_2(x))q(x)dx$$
  $$= \int_{-1}^{1} (\alpha p_1(x)q(x) + \beta p_2(x)q(x))d(x)$$
  $$= \int_{-1}^{1} \alpha p_1(x)q(x)dx + \int_{-1}^{1} \beta p_2(x)q(x)dx$$
  $$= \alpha \int_{-1}^{1} p_1(x)q(x)dx + \beta \int_{-1}^{1} p_2(x)q(x)dx$$
  $$= \alpha K(p_1, q) + \beta K(p_2, q).$$

- **Conjugate symmetry.**

  $$K(q, p) = \int_{-1}^{1} q(x)p(x)dx$$
  $$= \int_{-1}^{1} \overline{q(x)p(x)}dx$$
  $$= \int_{-1}^{1} \overline{p(x)q(x)}dx$$
  $$= \int_{-1}^{1} \overline{p(x)q(x)}dx$$
  $$= \overline{K(p, q)}.$$

- **Positive definiteness.** Note that $p(x)^2 \geq 0 \forall p \in \mathcal{P}_3, x \in [-1, 1]$. Therefore,

  $$K(p, p) = \int_{-1}^{1} p(x)p(x)dx = \int_{-1}^{1} p(x)^2dx \geq 0.$$

  Moreover, if we have $K(p, p) = \int_{-1}^{1} p(x)^2dx = 0$, since $p(x)^2$ is non-negative, we necessarily need $p(x)$ to be identically 0.

(b) First of all, recall that $1, x, x^2, x^3$ form a basis for $\mathcal{P}_3$ (cf. HW 1, Problem 8(a)). We will
use this basis to construct an orthonormal basis using Gram–Schmidt orthogonalization.

\[ \tilde{p}_0(x) = 1, \]
\[ p_0(x) = \frac{1}{\|1\|} = \frac{1}{\sqrt{2}} \]

\[ \tilde{p}_1(x) = x - \left( \int_{-1}^{1} \frac{x}{\sqrt{2}} \, dx \right) \cdot \frac{1}{\sqrt{2}} = x, \]
\[ p_1(x) = \frac{x}{\|x\|} = \frac{\sqrt{3}}{2} x \]

\[ \tilde{p}_2(x) = x^2 - \left( \int_{-1}^{1} \frac{3}{2} x^3 \, dx \right) \cdot \frac{3}{2} - \left( \int_{-1}^{1} \frac{x^2}{\sqrt{2}} \, dx \right) \cdot \frac{1}{\sqrt{2}} = x^2 - \frac{1}{3}, \]
\[ p_2(x) = \frac{x^2 - 1/3}{\|x^2 - 1/3\|} = \sqrt{\frac{45}{8}} \left( x^2 - \frac{1}{3} \right) \]

\[ \tilde{p}_3(x) = x^3 - \left( \int_{-1}^{1} \frac{45}{8} x^5 \, dx \right) \cdot \sqrt{\frac{45}{8}} \left( x^2 - \frac{1}{3} \right) - \left( \int_{-1}^{1} \frac{3}{2} x^4 \, dx \right) \cdot \frac{3}{2} x \]
\[ \quad - \left( \int_{-1}^{1} \frac{x^3}{\sqrt{2}} \, dx \right) \cdot \frac{1}{\sqrt{2}} \]
\[ = x^3 - \frac{3}{5} x, \]
\[ p_3(x) = \frac{x^3 - (3/5)x}{\|x^3 - (3/5)x\|} = \sqrt{\frac{175}{8}} \left( x^3 - \frac{3}{5} x \right). \]

4. Bessel’s inequality. Suppose that the columns of \( U \in \mathbb{R}^{n \times k} \) are orthonormal. Show that \( \|U'x\| \leq \|x\| \).

**Solution:** Let \( u_1, u_2, \ldots, u_k \) denote the columns of \( U \), and let

\[ \tilde{x} := \sum_{j=1}^{k} (u_j'x)u_j = UU'x \]
be the projection of \( x \) onto \( \mathcal{R}(U) \). Then

\[
0 \leq \|x - \tilde{x}\|^2 \\
= \|(I - UU')x\|^2 \\
= x'(I - UU')(I - UU')x \\
= x'(I - UU')x \\
= x'x - x'UU'x \\
= \|x\|^2 - \|U'x\|^2.
\]

5. **Householder reflections.** A **Householder matrix** is defined as

\[
Q = I - 2uu'
\]

for a unit vector \( u \in \mathbb{R}^n \).

(a) Show that \( Q \) is orthogonal.

(b) Show that \( Qu = -u \) and that \( Qv = v \) for every \( v \perp u \). Thus, the linear transformation \( y = Qx \) reflects \( x \) through the hyperplane with normal vector \( u \).

(c) Given \( y \), find \( x \) such that \( y = Qx \).

(d) Show that \( \det(Q) = -1 \).

(e) Given nonzero vectors \( x \) and \( y \), find a unit vector \( u \) such that

\[
Qx = (I - 2uu')x \in \text{span}(y),
\]

in terms of \( x \) and \( y \).

**Solution:**

(a) Consider

\[
Q'Q = (I - 2uu')(I - 2uu') \\
= (I - 2uu')(I - 2uu') \\
= I - 2uu' - 2uu' + 4uu'uu' \\
= I - 4uu' + 4uu' \\
= I.
\]

(b) For the unit vector \( u \),

\[
Qu = (I - 2uu')u \\
= u - 2uu'u \\
= u - 2u \\
= -u.
\]
For any \( v \perp u \),
\[
Qv = (I - 2uu')v \\
= v - 2uu'v \\
= v - 0 \\
= v.
\]

Any vector \( x \in \mathbb{R}^n \) can be written as \( x = (u'x)u + v \), where \( u'v = u'(x - (u'x)u) = 0 \), that is, \( v \perp u \). Hence, \( Qx = -(u'x)u + v \), which can be interpreted as the reflection of \( x \) through a hyperplane with normal vector \( u \).

(c) Since \( Q \) is symmetric and orthogonal, \( Q^{-1} = Q' = Q \). Hence, \( x = Q^{-1}y = Qy \), which is the reflection back from \( y \).

(d) Since \( \det(I - AB) = \det(I - BA) \), \( \det(Q) = \det(I - 2uu') = \det(1 - 2u'u) = -1 \).

(e) The question is asking to find the vector \( u \) such that \( x \) and \( \alpha y \) for some \( \alpha \) are reflections of each other through the plane \( \{ z : u'z = 0 \} \). Since \( Q \) is isometric, \( ||x|| = ||Qx|| = ||\alpha y|| \), which implies that \( \alpha = \pm ||x||/||y|| \). We will consider the case \( Qx = (||x||/||y||)y = \tilde{x} \), but the other case \( Qx = -(||x||/||y||)y = -\tilde{x} \) works equally well and provides an alternative answer. Now note that if the hyperplane \( \{ z : u'z = 0 \} \) reflects \( x \) to \( \tilde{x} \) and vice versa, then \( x - \tilde{x} \) is normal to the hyperplane, or equivalently,
\[
u = \frac{x - \tilde{x}}{||x - \tilde{x}||} = \frac{x - (||x||/||y||)y}{||x - (||x||/||y||)y||} = \frac{y||x - ||x||y}{(||y||x - ||x||y)}
\]
is the desired unit normal vector. Alternatively, if we take \(-\tilde{x}\) as the reflection of \( x \), then
\[
u = \frac{x + \tilde{x}}{||x + \tilde{x}||} = \frac{||y||x + ||x||y}{(||y||x + ||x||y)}.
\]

6. **Projection matrices.** A symmetric matrix \( P = P' \in \mathbb{R}^{n \times n} \) is said to be a **projection matrix** if \( P = P^2 \).

(a) Show that if \( P \) is a projection matrix, then so is \( I - P \).

(b) Suppose that the columns of \( U \in \mathbb{R}^{n \times k} \) are orthonormal. Show that \( UU' \) is a projection matrix.

(c) Suppose that \( A \in \mathbb{R}^{n \times k} \) is full-rank with \( k \leq n \). Show that \( A(A'A)^{-1}A' \) is a projection matrix.

(d) The point \( y \in S \subseteq \mathbb{R}^n \) closest to \( x \in \mathbb{R}^n \) is said to be the **orthogonal projection** (or **projection** in short) of \( x \) onto \( S \). Show that if \( P \) is a projection matrix, then \( y = Px \) is the projection of \( x \) onto \( \mathcal{R}(P) \).

(e) Let \( u \) be a unit vector. Find the projection matrix \( P \) such that \( y = Px \) is the projection of \( x \) onto \( \text{span}(u) \).

**Solution:**

(a) Note that \( (I - P)' = I - P' = I - P \) and so \( I - P \) is symmetric. Also, \( (I - P)^2 = (I - P)(I - P) = I - P - P + P^2 = I - 2P + P = I - P \). Hence, \( I - P \) is symmetric and \( (I - P)^2 = (I - P) \), thus it is a projection matrix.
(b) Note that \((UU')' = (U')'U = UU'\) and
\[
(UU')^2 = UU'UU' = UIU' = UU'.
\]

(c) First of all, \((A'A)\) is invertible for a full-rank tall matrix \(A\). For symmetry, consider
\[
(A(A'A)^{-1}A')' = (A')'(A'A)^{-1}'A'
= A((A'A)^{-1})'A'
= A((A'A)^{-1})A'
= A(A'A)^{-1}A'.
\]
Also, consider
\[
(A(A'A)^{-1}A')^2 = A(A'A)^{-1}(A'A)(A'A)^{-1}A'
= A(A'A)^{-1}A'.
\]

(d) It suffices to show that \(\|x - v\|\) is minimized over all \(v \in \mathcal{R}(P)\) by \(v^* = Px\). For any \(v \in \mathcal{R}(P)\),
\[
\|x - v\|^2 = \|x - Px + Px - v\|^2
= \|x - Px\|^2 + \|Px - v\|^2 + 2(x - Px)'(Px - v)
= \|x - Px\|^2 + \|Px - v\|^2 + 2(x' - x'P)(Px - v)
= \|x - Px\|^2 + \|Px - v\|^2 + 2(x'Px - x'v - x'P^2x + x'Pv)
\]
\[
(1) \quad \leq \|x - Px\|^2 + \|Px - v\|^2 + 2(x'Px - x'v - x'Px + x'v)
= \|x - Px\|^2 + \|Px - v\|^2
\]
\[
(2) \quad \leq \|x - Px\|^2,
\]
where (1) follows using \(P^2 = P\) and \(Pv = v \forall v \in \mathcal{R}(P)\). To achieve equality in (2), we need \(\|Px - v\| = 0 \Rightarrow v = Px\). Thus, \(\arg \min_{v \in \mathcal{R}(P)} \|x - v\| = Px\).

(e) Consider \(P = uu'\). Since \(P = P'\) and \(P = P^2\), \(P\) is a projection matrix. Since \(\mathcal{R}(P) = \text{span}(u)\), by part (d) \(Px\) is the projection of \(x\) onto \(\text{span}(u)\). This can be also directly verified since \((u'x)u = uu'x = Px\) is the component of \(x\) in the direction of the unit vector \(u\).

7. Reflection and projection with an affine hyperplane. Let \(a\) be a nonzero vector in \(\mathbb{R}^n\), \(b \in \mathbb{R}\), and
\[
\mathcal{A} = \{x \in \mathbb{R}^n : a'x = b\}.
\]
be an affine hyperplane, namely, a shifted version of the hyperplane \(\mathcal{H} = \{x : a'x = 0\}\) by \(b\), with the same normal vector \(a\).

(a) Find the projection of the zero vector \(0\) onto \(\mathcal{A}\).
(b) Find the reflection of 0 through \( A \).
(c) Find the projection of \( x \) onto \( A \).
(d) Find the reflection of \( x \) through \( A \).

**Solution:** If we have an affine hyperplane \( A \), \( p_A(x) \) for any \( x \) will be such that the vector \( p_A(x) - x \) is perpendicular to the plane \( A \). This can be seen by elementary geometric arguments—there always exists a point \( y \in A \) such that \( y - x \perp A \). If \( p_A(x) = z \neq y \), then the vectors \( y - x, z - y, x - z \) form a right-angled triangle with \( z - x \) as the hypotenuse and we have \( \| x - y \| \leq \| x - z \| \) contradicting the minimality in distance from \( x \), that is implicit in a projection. Similarly, the reflection of \( x \), \( r_A(x) \) is equidistant from \( p_A(x) \) as \( x \), but in the exact opposite direction.

(a) The vector \( (p_A(0) - 0) \) must be in the direction of \( a \). Thus, \( p_A(0) = \alpha a \). But, we also know that \( p_A(0) \) lies on the affine hyperplane \( A \). Thus, \( a'p_A(0) = b \). Putting these two together, we get that \( a'\alpha a = b \implies \alpha = \frac{b}{a'} \). Thus, \( p_A(0) = \frac{b}{a'} a \).

(b) Through geometric arguments, we can see that the reflection of 0 through \( A \) will be such that the vector \( r_A(0) - p_A(0) = (p_A(0) - 0) \). Thus \( r_A = 2p_A(0) = 2\frac{b}{a'} a \).

(c) We follow the same reasoning as in (a) to argue that \( (p_A(x) - x) \) must be in the direction of \( a \), and \( p_A(x) \) must lie on the affine hyperplane \( A \). Thus,

\[
p_A(x) - x = \alpha a
\]

and

\[
a'p_A(x) = b,
\]

which implies that \( a'(\alpha a + x) = b \), or equivalently, \( \alpha = \frac{b - a'x}{a'a} \). Therefore,

\[
p_A(x) = x - \frac{1}{\|a\|^2}(a'x - b)a.
\]

(d) Again, we use the same reasoning as in (b) to argue that the reflection \( r_A(x) \) of \( x \) must satisfy \( r_A(x) - p_A(x) = p_A(x) - x \), which implies that

\[
r_A(x) = 2p_A(x) - x = x - 2\frac{1}{\|a\|^2}(a'x - b)a.
\]