MATH 217 - LINEAR ALGEBRA
HOMEWORK 9, SOLUTIONS

Part A (10 points)

Solve the following problems from the book:
Section 5.4: #8;
Section 5.5: #11, 20, 23, 30.

Solution.

5.4.8. (a) We are given that $L(\vec{x}) = A\vec{x}$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ (so A is $m \times n$) with $\ker(L) = \{\vec{0}\}$. The *pseudoinverse* L^+ if L is defined by letting $L^+(\vec{y})$ be the least-squares solution to $L(\vec{x}) = \vec{y}$ (or equivalently, $A\vec{x} = \vec{y}$).

Since $\ker(L) = \ker(A) = \{\vec{0}\}$, $A^T A$ is invertible and the unique least-squares solution to $A\vec{x} = \vec{y}$ is

$$\vec{x}^* = (A^T A)^{-1} A^T \vec{y},$$

by Theorem 5.4.6. Thus, for any $\vec{y}$,

$$L^+(\vec{y}) = (A^T A)^{-1} A^T \vec{y},$$

showing that L^+ is a matrix transformation (and thus linear) with standard matrix $A^+ = (A^T A)^{-1} A^T$.

(b) If L is invertible, then A is an invertible matrix and so $A\vec{x} = \vec{y}$ has a unique solution $\vec{x} = A^{-1}\vec{y}$ for every $\vec{y}$, and consequently, the least-squares solution is also $L^+(\vec{y}) = A^{-1}\vec{y}$, for every $\vec{y}$. Thus, $L^+ = L^{-1}$ in this case.

(c) For any $\vec{x} \in \mathbb{R}^n$, if we set $\vec{y} = L(\vec{x})$, then $L(\vec{x}') = \vec{y}$ has a unique solution, namely $\vec{x}' = \vec{x}$, so this is also the least-squares solution and thus $L^+(L(\vec{x})) = \vec{x}$.

(d) For any $\vec{y} \in \mathbb{R}^m$,

$$L(L^+(\vec{y})) = A((A^T A)^{-1} A^T \vec{y}) = A(A^T A)^{-1} A^T \vec{y},$$

which by Theorem 5.4.7 is exactly the orthogonal projection of $\vec{y}$ onto $\mathrm{im}(A) = \mathrm{im}(L)$.

(e) If L has standard matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

then by the formula we found in part (a), L^+ is the matrix transformation with standard matrix

$$\begin{aligned}
A^+ &= (A^TA)^{-1}A^T \\
&= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.
\end{aligned}$$

5.5.11. Given $f(t) = \cos(t)$ and $g(t) = \cos(t + \delta)$ in $C([-\pi, \pi])$, where $0 \leq \delta \leq \pi$, we have that

$$\begin{aligned}
\langle f, g \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)g(t)\, dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(t)\cos(t+\delta)\, dt \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(t)[\cos(t)\cos(\delta) - \sin(t)\sin(\delta)]\, dt \\
&= \frac{1}{\pi} \left[\cos(\delta) \int_{-\pi}^{\pi} \cos^2(t)\, dt - \sin(\delta) \int_{-\pi}^{\pi} \cos(t)\sin(t)\, dt \right] \\
&= \frac{1}{\pi}[\cos(\delta)\pi - 0] \\
&= \cos(\delta),
\end{aligned}$$

and

$$\|f\|^2 = \langle f, f \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(t)\, dt = 1$$

$$\|g\|^2 = \langle g, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(t+\delta)\, dt = 1,$$

so

$$\measuredangle(f, g) = \arccos \frac{\langle f, g \rangle}{\|f\|\|g\|} = \arccos \frac{\cos(\delta)}{1} = \delta.$$

5.5.20. We are given the inner product

$$\langle \vec{v}, \vec{w} \rangle = \vec{v}^T \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \vec{w}$$

on $\mathbb{R}^2$.

(a) Let $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ and suppose that $\left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle = 0$, that is:

$$0 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = x + 2y.$$

Thus, the vectors $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ which are orthogonal to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with respect to *this* inner product are exactly those lying on the line $x + 2y = 0$.

(b) To find an orthonormal basis for $\mathbb{R}^2$ with respect to this inner product, consider the vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. These vectors are orthogonal by part (a), since the latter vector lies on $x + 2y = 0$. It remains to normalize them:

$$\left\| \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\| = \sqrt{\left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle} = \sqrt{\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \sqrt{\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}} = 1$$

and

$$\left\| \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\| = \sqrt{\left\langle \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\rangle} = \sqrt{\begin{bmatrix} 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix}} = \sqrt{\begin{bmatrix} 2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -4 \end{bmatrix}} = 2.$$

Thus,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \frac{1}{2} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

is an orthonormal basis.

5.5.23. We are given the inner product

$$\langle f, g \rangle = \frac{1}{2}(f(0)g(0) + f(1)g(1))$$

on P_1, the space of polynomials of degree ≤ 1. To find an orthonormal basis, we begin with the basis $(1, x)$ and perform the Gram-Schmidt process using the above inner product.

$$\|1\| = \sqrt{\frac{1}{2}(1 \cdot 1 + 1 \cdot 1)} = 1$$

so we can let $u_1 = 1$. Next, let

$$w_2 = x - \langle x, u_1 \rangle u_1 = x - \langle x, 1 \rangle 1 = x - \sqrt{\frac{1}{2}(0 \cdot 1 + 1 \cdot 1)} = x - \sqrt{1/2}.$$

Then,

$$\begin{aligned} \|w_2\| &= \sqrt{\frac{1}{2}((-\sqrt{1/2})^2 + (1 - \sqrt{1/2})^2)} \\ &= \sqrt{\frac{1}{2}(1/2 + 1 - 2\sqrt{1/2} + 1/2)} \\ &= \sqrt{1 - \sqrt{1/2}}, \end{aligned}$$

and so letting

$$u_1 = \frac{1}{\|w_2\|} w_2 = \frac{1}{\sqrt{1 - \sqrt{1/2}}} \left(x - \sqrt{1/2} \right),$$

we have that (u_1, u_2) is an orthonormal basis for P_1 with this inner product.

5.5.30. Recall that an ellipse E centered at the origin can be represented as all points (x, y) such that

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

for some $a, b > 0$ (the half-widths of the axes lying along the x and y-axes, respectively). See Demsos here.

Define $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^2$ as follows:

$$\left\langle \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right\rangle = \frac{x_1 x_2}{a^2} + \frac{y_1 y_2}{b^2}.$$

It is immediate then from this definition that E is exactly all those $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ such that $\|\vec{x}\|^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ (and thus $\|\vec{x}\|$) is equal to 1. It remains to verify that $\langle \cdot, \cdot \rangle$ is, in fact, an inner product.

Let $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \in \mathbb{R}^2$ and $k \in \mathbb{R}$. For symmetry, observe that

$$\left\langle \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right\rangle = \frac{x_1 x_2}{a^2} + \frac{y_1 y_2}{b^2} = \frac{x_2 x_1}{a^2} + \frac{y_2 y_1}{b^2} = \left\langle \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \right\rangle.$$

For bilinearity,

$$\begin{aligned} \left\langle \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \right\rangle &= \frac{(x_1 + x_2)x_3}{a^2} + \frac{(y_1 + y_2)y_3}{b^2} \\ &= \frac{x_1 x_2 + x_2 x_3}{a^2} + \frac{y_1 y_3 + y_2 y_3}{b^2} \\ &= \frac{x_1 x_3}{a^2} + \frac{y_1 y_3}{b^2} + \frac{x_2 x_3}{a^2} + \frac{y_2 y_3}{b^2} \\ &= \left\langle \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \right\rangle + \left\langle \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \right\rangle, \end{aligned}$$

and

$$\left\langle k \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} kx_1 \\ ky_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right\rangle = \frac{kx_1 x_2}{a^2} + \frac{ky_1 y_2}{b^2} = k\left(\frac{x_1 x_2}{a^2} + \frac{y_1 y_2}{b^2}\right) = k\left\langle \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right\rangle.$$

This show linearity in the first coordinate, linearity in the second coordinate follows by symmetry. Lastly, for positive definiteness, if $\begin{bmatrix} x \\ y \end{bmatrix} \neq \vec{0}$, then either $x \neq 0$ or $y \neq 0$, and so

$$\left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle = \frac{x^2}{a^2} + \frac{y^2}{b^2} > 0,$$

since $a, b > 0$. Thus, $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{R}^2$.

Part B (25 points)

Problem 1. State whether each of the following is TRUE or FALSE for an arbitrary $m \times n$ matrix A, and provide either a short proof or counterexample for your claim.

(a) $A^\top A = AA^\top$.
(b) $\operatorname{rank}(A) = \operatorname{rank}(A^\top)$.
(c) $\dim(\ker(A)) = \dim(\ker(A^\top))$.
(d) If $\ker(A) = \{\vec{0}\}$, then $A^\top A$ is invertible.
(e) If $\ker(A) = \{\vec{0}\}$, then $AA^\top$ is invertible.

Solution.

(a) FALSE: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, then

$$A^T A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 13 \end{bmatrix},$$

but

$$AA^T = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 6 & 9 \end{bmatrix} \neq A^T A.$$

(b) TRUE: By Rank-Nullity applied to the $n \times m$ matrix A^T, we have that

$$m = \dim(\operatorname{im}(A^T)) + \dim(\ker(A^T)) = \operatorname{rank}(A^T) + \dim(\ker(A^T)).$$

By Theorem 5.4.1, $\ker(A^T) = (\operatorname{im}(A))^\perp$, while by a Worksheet problem, for any subspace V of $\mathbb{R}^m$, $\dim(V) + \dim(V^\perp) = m$, so

$$\operatorname{rank}(A) + \dim(\ker(A^T)) = \dim(\operatorname{im}(A)) + \dim(\operatorname{im}(A)^\perp) = m.$$

Putting this together, we have that

$$\operatorname{rank}(A^T) + \dim(\ker(A^T)) = m = \operatorname{rank}(A) + \dim(\ker(A^T)),$$

and so

$$\operatorname{rank}(A^T) = \operatorname{rank}(A).$$

(c) FALSE: Consider

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then, $\dim(\ker(A)) = 0$, but

$$A^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

and so $\dim(\ker(A^T)) = 1 \neq \dim(\ker(A))$.

(d) TRUE: By Theorem 5.4.2, $\ker(A) = \ker(A^T A)$, so if $\ker(A) = \{\vec{0}\}$, then $\ker(A^T A) = \{\vec{0}\}$ as well, and since $A^T A$ is $n \times n$, it must be invertible by Theorem 3.1.7.

(e) FALSE: Consider the example from part (c) above. It was shown there that $\ker(A) = \{\vec{0}\}$, but

$$AA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

which is clearly not invertible.

Problem 2. Recall the definitions of the sum $V+W$ and orthogonal complement $V^\perp$ of subspaces V, W of $\mathbb{R}^n$ from previous homework.

(a) Prove that for any subspace V of $\mathbb{R}^n$, $\mathbb{R}^n = V + V^\perp$.

(b) Prove that for any $m \times n$ matrix A, $\mathbb{R}^m = \operatorname{im}(A) + \ker(A^\top)$ and $\mathbb{R}^n = \operatorname{im}(A^\top) + \ker(A)$.

Thinking of A as the standard matrix of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, this shows that we can "decompose" the domain space $\mathbb{R}^n$ as the sum of the *image* (or *column space*) $\operatorname{im}(A)$ and the *cokernel* $\ker(A^\top)$ of A, and the target space $\mathbb{R}^m$ as the sum of the *row space* $\operatorname{im}(A^\top)$ and the *kernel* $\ker(A)$.

Solution.

(a) We need to show that every $\vec{x} \in \mathbb{R}^n$ can be expressed as $\vec{x} = \vec{v} + \vec{v}^\perp$, where $\vec{v} \in V$ and $\vec{v}^\perp \in V^\perp$. If we let $\vec{v} = \operatorname{proj}_V(\vec{x})$ and $\vec{v}^\perp = \vec{x} - \operatorname{proj}_V(\vec{x})$, then by definition of the orthogonal projection $\operatorname{proj}_V(\vec{x})$ of $\vec{x}$ onto V, $\vec{v} \in V$ and $\vec{v}^\perp \in V^\perp$. Moreover, it is clear that

$$\vec{x} = \operatorname{proj}_V(\vec{x}) + (\vec{x} - \operatorname{proj}_V(\vec{x})) = \vec{v} + \vec{v}^\perp \in V + V^\perp,$$

which proves the claim.

(b) By Theorem 5.4.1, $\ker(A^T) = (\operatorname{im}(A))^\perp$, so by part (a) applied to $V = \operatorname{im}(A)$ in $\mathbb{R}^m$, we have that

$$\mathbb{R}^m = \operatorname{im}(A) + \operatorname{im}(A)^\perp = \operatorname{im}(A) + \ker(A^\top).$$

By Theorem 5.4.1 applied to the matrix A^T, $\ker(A) = \ker((A^T)^T) = \operatorname{im}(A^T)^\perp$, so again by part (a), this time applied to $V = \operatorname{im}(A^T)$ in $\mathbb{R}^n$, we have that

$$\mathbb{R}^n = \operatorname{im}(A^T) + \operatorname{im}(A^T)^\perp = \operatorname{im}(A^T) + \ker(A),$$

which proves the claim.

Problem 3. Let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space V, with norm $\| \cdot \|$ defined by $\|v\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$ for all $v \in V$. Prove each of the following:

(a) (*Pythagorean Theorem*) For all $v, w \in V$, the vectors v and w are orthogonal if and only if

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

(b) (*Cauchy–Schwarz Inequality*) For all $v, w \in V$,

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

(Hint: Write $v = \operatorname{proj}_w(v) + (v - \operatorname{proj}_w(v))$ and apply (a) to obtain $\|\operatorname{proj}_w(v)\| \leq \|v\|$.)

(c) (*Triangle Inequality*) For all $v, w \in V$,

$$\|v + w\| \leq \|v\| + \|w\|.$$

(d) (*Polarization Identity*) For all $v, w \in V$,

$$\langle v, w\rangle = \tfrac{1}{4}\|v+w\|^2 - \tfrac{1}{4}\|v-w\|^2.$$

Solution. Throughout, let $v, w \in V$.

(a) We can expand the left-hand side of the desired equation and use properties of the inner product:

$$\begin{aligned}\|v+w\|^2 &= \langle v+w, v+w\rangle \\ &= \langle v, v+w\rangle + \langle w, v+w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\ &= \|v\|^2 + 2\langle v, w\rangle + \|w\|^2.\end{aligned}$$

But then, $\|v\|^2 + 2\langle v, w\rangle + \|w\|^2 = \|v\|^2 + \|w\|^2$ if and only if $\langle v, w\rangle = 0$, that is, if and only if v and w are orthogonal.

(b) We'll split this into two cases: If $w = \vec{0}$, then

$$|\langle v, w\rangle| = |\langle v, \vec{0}\rangle| = |\langle v, 0\cdot\vec{0}\rangle| = |0\langle v, \vec{0}\rangle| = 0 \leq \|v\|\|w\|,$$

proving the inequality in this case.

Now we can assume that $w \neq \vec{0}$. Then, we can write

$$v = \mathrm{proj}_w(v) + (v - \mathrm{proj}_w(v)),$$

where $\mathrm{proj}_w(v) = \frac{\langle v,w\rangle}{\|w\|^2}w$. Observe that

$$\begin{aligned}\langle \mathrm{proj}_w(v), v - \mathrm{proj}_w(v)\rangle &= \left\langle \frac{\langle v, w\rangle}{\|w\|^2}w, v - \frac{\langle v, w\rangle}{\|w\|^2}w\right\rangle \\ &= \left\langle \frac{\langle v, w\rangle}{\|w\|^2}w, v\right\rangle - \left\langle \frac{\langle v, w\rangle}{\|w\|^2}w, \frac{\langle v, w\rangle}{\|w\|^2}w\right\rangle \\ &= \frac{\langle v, w\rangle^2}{\|w\|^2} - \frac{\langle v, w\rangle^2}{\|w\|^4}\|w\|^2 = 0,\end{aligned}$$

so $\mathrm{proj}_w(v)$ and $v - \mathrm{proj}_w(v)$ are orthogonal. We can thus apply part (a) to $v = \mathrm{proj}_w(v) + (v - \mathrm{proj}_w(v))$ to get

$$\|v\|^2 = \|\mathrm{proj}_w(v) + (v - \mathrm{proj}_w(v))\|^2 = \|\mathrm{proj}_w(v)\|^2 + \|v - \mathrm{proj}_w(v)\|^2.$$

Since $\|v - \mathrm{proj}_w(v)\|^2 \geq 0$, we have that

$$\|v\|^2 \geq \|\mathrm{proj}_w(v)\|^2$$

and so

$$\|v\| \geq \|\mathrm{proj}_w(v)\| = \left\|\frac{\langle v, w\rangle}{\|w\|^2}w\right\| = |\langle v, w\rangle|\frac{\|w\|}{\|w\|^2} = \frac{|\langle v, w\rangle|}{\|w\|}.$$

Multiplying both sides of this inequality by $\|w\| > 0$, we have that

$$|\langle v, w\rangle| \leq \|v\|\|w\|,$$

as desired.

(c) As in the proof of (a),

$$\|v+w\|^2 = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2,$$

but by (b),

$$\langle v, w\rangle \leq |\langle v, w\rangle| \leq \|v\|\|w\|,$$

so putting these together we have that

$$\|v+w\|^2 = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

Taking square roots, we obtain

$$\|v+w\| \leq \|v\| + \|w\|.$$

(d) We can expand the right-hand of the desire equation to obtain

$$\begin{aligned}
\frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2 &= \frac{1}{4}\langle v+w, v+w\rangle - \frac{1}{4}\langle v-w, v-w\rangle \\
&= \frac{1}{4}(\|v\|^2 + 2\langle v, w\rangle + \|w\|^2) - \frac{1}{4}(\|v\|^2 - 2\langle v, w\rangle + \|w\|^2) \\
&= \frac{1}{2}\langle v, w\rangle + \frac{1}{2}\langle v, w\rangle \\
&= \langle v, w\rangle,
\end{aligned}$$

which completes the proof.

Problem 4. Consider the vector space $C([-\pi, \pi])$ of all continuous functions from $[-\pi, \pi]$ to $\mathbb{R}$, and define $\langle \cdot, \cdot \rangle$ on $C([-\pi, \pi])$ by

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\, dx$$

for all $f, g \in C([-\pi, \pi])$. In parts (b) and (c) below, you may use technology to help evaluate any integrals, but you must set them up explicitly first.

(a) Prove that $\langle \cdot, \cdot \rangle$ is an inner product on $C([-\pi, \pi])$.[1]

(b) Recall that Theorems 5.5.4 and 5.5.5 of the textbook show that

$$\frac{1}{\sqrt{2}},\ \sin(t),\ \cos(t),\ \sin(2t),\ \cos(2t), \ldots,\ \sin(nt),\ \cos(nt)$$

forms an orthonormal basis for its span T_n in $C([-\pi, \pi])$, and for any $f \in C([-\pi, \pi])$, the orthogonal projection of f onto T_n is given by the nth order Fourier approximation of f .

  (i) For each $n \in \mathbb{N}$, find the nth order Fourier approximation of the function $h(x) = x^2$.
  (ii) (**Recreational**) Fill in the coefficients (the "?"s) here: www.desmos.com/calculator/nux4fmn7sv. Press ▶ to see this approximation in action.

(c) Consider the functions $f(x) = 1$, $g(x) = x$, and $h(x) = x^2$ in $C([-\pi, \pi])$, and let V the be subspace spanned by them.

  (i) Find an orthonormal basis for V with respect to the inner product $\langle \cdot, \cdot \rangle$ above.

[1] You may use without proof the following fact from calculus: if $f : [a, b] \to \mathbb{R}$ is a continuous function that is nonnegative everywhere on the interval $[a, b]$, and strictly positive somewhere on $[a, b]$, then $\int_a^b f(x)dx > 0$. (Draw a picture to convince yourself of this!)

(ii) What element of V is nearest to the function $k(x) = x^3$ with respect to the norm induced by $\langle \cdot, \cdot \rangle$? And what is its distance to k?

Solution.

(a) Let $f, g, h \in C \in C([-\pi, \pi])$ and $c \in \mathbb{R}$. To see that $\langle \cdot, \cdot \rangle$ is symmetric, observe that

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)\,dx = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x)f(x)\,dx = \langle g, f \rangle.$$

For bilinearity (note that by symmetry it suffices to prove linearity in the first coordinate),

$$\begin{aligned}\langle f+g, h \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} (f(x)+g(x))h(x)\,dx \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} (f(x)h(x)+g(x)h(x))\,dx \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)h(x)\,dx + \frac{1}{\pi} \int_{-\pi}^{\pi} g(x)h(x)\,dx \\ &= \langle f, h \rangle + \langle g, h \rangle,\end{aligned}$$

and

$$\langle cf, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} cf(x)g(x)\,dx = \frac{c}{\pi} \int_{-\pi}^{\pi} f(x)g(x)\,dx = c\langle f, g \rangle.$$

For positive definiteness, suppose that $f \neq 0$. Recall that in this vector space, this means that there is some $x_0 \in [-\pi, \pi]$ such that $f(x_0) \neq 0$. Then, $f(x)^2 \geq 0$ for all $x \in [-\pi, \pi]$, $f(x_0)^2 > 0$, and so since f is continuous, there is some interval $[a, b]$ containing x_0, with $-\pi \leq a < b \leq \pi$, on which $f(x)^2 > 0$. Then, by properties of the integral and the fact in the footnote, we have that

$$\langle f, f \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2\,dx \geq \frac{1}{\pi} \int_{a}^{b} f(x)^2\,dx > 0.$$

This completes the proof.

(b) (i) We follow the notation (a_0, b_k, and c_k) in Theorem 5.5.5.

$$a_0 = \left\langle x^2, \frac{1}{\sqrt{2}} \right\rangle = \frac{1}{\sqrt{2}\pi} \int_{-\pi}^{\pi} x^2\,dx = \frac{1}{\sqrt{2}\pi} \frac{x^3}{3}\bigg|_{-\pi}^{\pi} = \frac{\sqrt{2}\pi^2}{3}$$

For $k \in \mathbb{N}$,

$$b_k = \langle x^2, \sin(kx) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \sin(kx)\,dx = 0,$$

since $x^2 \sin(kx)$ is odd (x^2 is even, $\sin(kx)$ is odd) and the integral is evaluated over a symmetric intervals. For the c_k's,

$$c_k = \langle x^2, \cos(kx) \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \cos(kx)\,dx = \frac{1}{\pi} \frac{4\pi(-1)^k}{k^2} = \frac{4(-1)^k}{k^2}.$$

To evaluate this integral, you can either use integration by parts twice, or technology, together with the facts that $\sin(k\pi) = 0$ and $\cos(k\pi) = (-1)^k$ for $k \in \mathbb{N}$. Thus, the nth order Fourier approximation of $h(x) = x^2$ is:

$$\begin{aligned} h_n(x) &= \left(\frac{\sqrt{2}\pi^2}{3}\right)\frac{1}{\sqrt{2}} + 0\sin(t) - 4\cos(t) + 0\sin(2t) + \frac{4}{4}\cos(2t) + \cdots + 0\sin(nt) + \frac{4(-1)^n}{n^2}\cos(nt) \\ &= \frac{\pi^2}{3} + \sum_{k=1}^{n} \frac{4(-1)^k}{k^2}\cos(kt). \end{aligned}$$

(ii)(**Recreational**) See here for the result: www.desmos.com/calculator/cwqspozgzc.

(c) (i) We perform the Gram-Schmidt process on the basis $(1, x, x^2)$ using the above inner product. Let $w_1(x) = 1$. Then,

$$\|w_1\| = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} 1\,dx} = \sqrt{\frac{2\pi}{\pi}} = \sqrt{2},$$

so we let

$$u_1(x) = \frac{1}{\|w_1\|}w_1 = \frac{1}{\sqrt{2}}.$$

Next, let

$$\begin{aligned} w_2(x) &= x - \langle x, u_1\rangle u_1 \\ &= x - \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{x}{\sqrt{2}}\,dx\right)\frac{1}{\sqrt{2}} \\ &= x, \end{aligned}$$

since $\int_{-\pi}^{\pi} x\,dx = 0$. Note that

$$\|w_2\| = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} x^2\,dx} = \sqrt{\frac{1}{\pi}\frac{2\pi^3}{3}} = \sqrt{\frac{2\pi^2}{3}},$$

so we let

$$u_2(x) = \frac{1}{\|w_2\|}w_2(x) = \sqrt{\frac{3}{2\pi^2}}x.$$

Lastly, we let

$$\begin{aligned} w_3(x) &= x^2 - \langle x^2, u_1\rangle u_1 - \langle x^2, u_2\rangle u_2 \\ &= x^2 - \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{x^2}{\sqrt{2}}\,dx\right)\frac{1}{\sqrt{2}} - \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \sqrt{\frac{3}{2\pi^2}}x^3\,dx\right)\sqrt{\frac{3}{2\pi^2}}x \\ &= x^2 - \frac{1}{2\pi}\int_{-\pi}^{\pi} x^2\,dx - 0 \\ &= x^2 - \frac{1}{2\pi}\frac{2\pi^3}{3} \\ &= x^2 - \frac{\pi^2}{3}, \end{aligned}$$

and since

$$\|w_3\| = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} (x^2 - \frac{\pi^2}{3})^2\,dx} = \sqrt{\frac{1}{\pi}\frac{8\pi^5}{45}} = \sqrt{\frac{8\pi^4}{45}},$$

we take

$$u_3(x) = \frac{1}{\|w_3\|} w_3(x) = \sqrt{\frac{45}{8\pi^4}} \left(x^2 - \frac{\pi^2}{3}\right).$$

Thus,

$$\left(\frac{1}{\sqrt{2}}, \quad \sqrt{\frac{3}{2\pi^2}} x, \quad \sqrt{\frac{45}{8\pi^4}} \left(x^2 - \frac{\pi^2}{3}\right) \right)$$

forms an orthonormal basis for this subspace.

(ii) We can find the nearest element to $k(x) = x^3$ to V, i.e., the orthogonal projection of $k(x)$ onto V, using the orthonormal basis from (i):

$$\operatorname{proj}_V(x^3) = \left\langle x^3, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}} + \left\langle x^3, \sqrt{\frac{3}{2\pi^2}} x \right\rangle \sqrt{\frac{3}{2\pi^2}} x + \left\langle x^3, \sqrt{\frac{45}{8\pi^4}} \left(x^2 - \frac{\pi^2}{3}\right) \right\rangle \sqrt{\frac{45}{8\pi^4}} \left(x^2 - \frac{\pi^2}{3}\right)$$

$$= \left(\frac{1}{\pi} \int_{-\pi}^{\pi} \frac{x^3}{\sqrt{2}}\, dx \right) \frac{1}{\sqrt{2}} + \left(\frac{1}{\pi} \int_{-\pi}^{\pi} \sqrt{\frac{3}{2\pi^2}} x^4\, dx \right) \sqrt{\frac{3}{2\pi^2}} x + \left(\frac{1}{\pi} \int_{-\pi}^{\pi} \sqrt{\frac{45}{8\pi^4}} \left(x^2 - \frac{\pi^2}{3}\right) x^3\, dx \right) \sqrt{\frac{45}{8\pi^4}} \left(x^2 - \frac{\pi^2}{3}\right).$$

Notice that the integrands in the first and third integral above are odd functions, so their integrals over the symmetric interval $[-\pi, \pi]$ are 0. This leaves

$$\begin{aligned} \operatorname{proj}_V(x^3) &= \left(\frac{1}{\pi} \int_{-\pi}^{\pi} \sqrt{\frac{3}{2\pi^2}} x^4\, dx \right) \sqrt{\frac{3}{2\pi^2}} x \\ &= \frac{3}{2\pi^3} \left(\int_{-\pi}^{\pi} x^4\, dx \right) x \\ &= \frac{3}{2\pi^3} \frac{2\pi^5}{5} x \\ &= \frac{3\pi^2}{5} x. \end{aligned}$$

The distance from $k(x) = x^3$ to V is then given by $\|x^3 - \operatorname{proj}_V(x^3)\|$:

$$\|x^3 - \operatorname{proj}_V(x^3)\| = \sqrt{\frac{1}{\pi} \int_{-\pi}^{\pi} \left(x^3 - \frac{3\pi^2}{5} x \right)^2 dx} = \sqrt{\frac{1}{\pi} \frac{8\pi^7}{175}} = \sqrt{\frac{8\pi^6}{175}} \approx 6.629.$$

Definition. Let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ be inner product spaces. A bijective linear map $T : V \to W$ is called an *isometry* if for all $x, y \in V$, we have $\langle x, y \rangle_V = \langle T(x), T(y) \rangle_W$. The inner product spaces V and W are said to be *isometric* if there exists an isometry from V to W.

Problem 5 (Recreational). Throughout this problem, assume $n \in \mathbb{N}$.

(a) Let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ be inner product spaces, and let $T : V \to W$ be a linear map such that $\langle x, y \rangle_V = \langle T(x), T(y) \rangle_W$ for all $x, y \in V$. Prove that T is injective.

(b) Let $\mathcal{U}$ be an orthonormal basis of the n-dimensional inner product space $(V, \langle \cdot, \cdot \rangle_V)$. Let $L_{\mathcal{U}}$ be the coordinate isomorphism from V to $\mathbb{R}^n$. Prove that $L_{\mathcal{U}}$ is an isometry from V to $\mathbb{R}^n$ with the dot product.

(c) Prove that every n-dimensional inner product space is isometric to $\mathbb{R}^n$ with the dot product.

(d) Prove that any two n-dimensional inner product spaces are isometric to each other.

Solution.

(a) Let $\vec{v}$ be a nonzero vector in V. Then $\langle T(\vec{v}), T(\vec{v})\rangle_W = \langle \vec{v}, \vec{v}\rangle \neq \vec{0}$ since $\langle \cdot, \cdot \rangle$ is positive-definite, so $\vec{v} \notin \ker(T)$. Thus $\ker(T) = \{\vec{0}\}$, so T is injective.

(b) To show that $L_{\mathcal{U}}$ is an isometry, we need to show that for all $x, y \in V$, $\langle x, y\rangle_V = [x]_{\mathcal{U}} \cdot [y]_{\mathcal{U}}$. Let $x, y \in V$, and write $x = \sum_{i=1}^n a_i u_i$ and $y = \sum_{i=1}^n b_i u_i$, where $\mathcal{U} = (u_1, \ldots, u_n)$. Then since $\mathcal{U}$ is orthonormal, we have

$$\langle x, y\rangle_V = \left\langle \sum_{i=1}^n a_i u_i, \sum_{j=1}^n b_j u_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \langle u_i, u_j\rangle = \sum_{i=1}^n a_i b_i = [x]_{\mathcal{U}} \cdot [y]_{\mathcal{U}}.$$

(c) Let $(V, \langle \cdot, \cdot \rangle)$ be an n-dimensional inner product space. By Gram-Schmidt, we can fix an orthonormal basis $\mathcal{U}$ of V. But then by part (b), $L_{\mathcal{U}}$ is an isometry from V to $\mathbb{R}^n$, so $(V, \langle \cdot, \cdot \rangle)$ is indeed isometric to $\mathbb{R}^n$ with the dot product.

(d) Let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ be n-dimensional inner product spaces. By part (c), both $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ are isometric to $\mathbb{R}^n$ with the dot product. But it is easy to see that inverses and compositions of isometries are isometries, so it follows that $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ are isometric to each other.