MATH 217 - W24 - LINEAR ALGEBRA
HOMEWORK 9, SOLUTIONS

Part A

Solve the following problems from Bretscher:
5.4: 27, 31.

Solution.

5.4.27 The least-squares solution(s) of the system $SA\vec{x} = S\vec{b}$ are the solutions of its normal equation equation, which is $(SA)^\top SA\vec{x} = (SA)^\top S\vec{b}$. We have:

$$\begin{aligned} (SA)^\top SA\vec{x} &= (SA)^\top S\vec{b} \\ \iff A^\top S^\top SA\vec{x} &= A^\top S^\top S\vec{b} \\ \iff A^\top A\vec{x} &= A^\top \vec{b} \end{aligned}$$

This last (equivalent) equation is the normal equation for the linear system $A\vec{x} = \vec{b}$. So $SA\vec{x} = S\vec{b}$ and $A\vec{x} = \vec{b}$ have the same least-squares solutions, and therefore the unique least-squares solution of $SA\vec{x} = S\vec{b}$ is $\begin{bmatrix} 7 \\ 11 \end{bmatrix}$.

5.4.31 The system of equations

$$\begin{aligned} c_0 + c_1(0) &= 3 \\ c_0 + c_1(1) &= 3 \\ c_0 + c_1(1) &= 6 \end{aligned}$$

is obviously inconsistent; in matrix form, it reads

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 6 \end{bmatrix}$$

The normal equation is

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 12 \\ 9 \end{bmatrix}$$

which has the solution

$$c_0 = 3, c_1 = 3/2$$

, and therefore the least-squares regression line is $f(x) = 3 + \frac{3}{2}x$.

5.5: 15, 23, 32(a, b, c, d).

Solution.

5.5.15 The proposed inner product is always bilinear, which can easily be shown by a computation, and is symmetric if and only if $b = c$, which can easily be shown by another computation.

The tricky part is positive definiteness. Given that $b = c$ as required above, we have $\|x\|^2 = x_1^2 + 2bx_1x_2 + dx_2^2$. Treating this as a quadratic in x_1, we find the discriminant is $4b^2x_2^2 - 4dx_2^2$. This quantity is negative whenever $x_2 \neq 0$ and $b^2 < d$. If we choose b and d so that this inequality holds, we have that the norm is zero only when x_2 is zero, which also, by the quadratic formula, implies that $x_1 = 0$. Thus we would have a positive definite bilinear form; i.e. an inner product. If $b^2 \geq d$, then there are we find at least one nontrivial solution to $\|x\| = 0$, and the bilinear form is therefore not positive definite.

5.5.23 Apply Gram-Schmidt to the basis $\{1, x\}$ for P_1. We find that $\langle 1, 1\rangle = 1$, and so we let $u_1 = 1$ in our orthonormal basis. Then we compute $\langle 1, x\rangle = 1/2$, and so take (our second, orthogonal but not normalized basis vector) $b_2 = x - 1/2$. Computing $\langle b_2, b_2\rangle = 1/2$, we then have that $u_2 = \sqrt{2}b_2 = \sqrt{2}x - \sqrt{2}/2$.

5.5.32 (a) In this case we are evaluating

$$\frac{1}{2}\int_{-1}^{1} t^{(n+m)}dt.$$

We use the power rule to find the antiderivative

$$\frac{1}{2(n+m+1)}t^{n+m+1}$$

evaluated on $[-1, 1]$, which gives

$$\frac{1}{2(n+m+1)}(1^{n+m+1} - (-1)^{n+m+1})$$

which is equal to zero if $n + m$ is odd, and is $\frac{1}{n+m+1}$ for $n + m$ even.

(b) Using (a), we are in the case that $n = m$, so $n + m = 2n$ which is even. Thus

$$\|t^n\| = \sqrt{\frac{1}{2n+1}}.$$

(c) We proceed bravely, using the formulae found in parts a and b repeatedly. Firstly observe that $\langle 1, 1\rangle = 1$, so $g_0(t) = 1$. Since $\langle t, 1\rangle = 0$, we need only normalize t to find g_1: $\langle t, t\rangle = 1/3$ so $g_1(t) = (\sqrt{3})t$.
Now we compute that $\langle t^2, g_0\rangle = 1/3$ while $\langle t^2, g_1\rangle = 0$. So then (before normalizing the length) $h_2(t) = t^2 - 1/3$. We find that $\langle t^2 - 1/3, t^2 - 1/3\rangle = 4/45$, and thus $g_2(t) = \sqrt{5}/2(3t^2 - 1)$.
Lastly, observe that $\langle t^3, g_0\rangle = 0$, $\langle t^3, g_1\rangle = \sqrt{3}/5$ and $\langle t^3, g_2\rangle = 0$. Therefore (before normalizing) $h_3 = t^3 - 3/5t$, $\langle h_3, h_3\rangle = 4/175$ so normalizing gives $g_3(t) = \sqrt{7}/2(5t^3 - 3t)$.

(d) Call the degree n Legendre polynomial $l_n(t)$. By a computation, we find $l_0(t) = 1$, $l_1(t) = t$, $l_2(t) = 1/2(3t^2 - 1)$, $l_3(t) = 1/2(5t^3 - 3t)$.

Part B

Problem 1. Consider the four points $(2, 4, 6), (1, 3, 2), (1, 1, 0)$ and $(1, 2, 3)$ in $\mathbb{R}^3$.

(a) Write a matrix equation that, *if it were consistent*, could be used to find the coefficients A, B, C in the equation of a plane of the form $z = Ax + By + C$ that contains all four points.

Solution. Using the four points to write a linear system, we get

$$\begin{aligned} 2A + 4B \quad +C &= 6 \\ A + 3B \quad +C &= 2 \\ A + B \quad +C &= 0 \\ A + 2B \quad +C &= 3 \end{aligned}$$

which corresponds to

$$\begin{bmatrix} 2 & 4 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ 0 \\ 3 \end{bmatrix}$$

(b) Show that the matrix equation from (a) is, in fact, inconsistent.

Solution. We use elementary row operations to put the augmented matrix in reduced row echelon form:

$$\left[\begin{array}{ccc|c} 2 & 4 & 1 & 6 \\ 1 & 3 & 1 & 2 \\ 1 & 1 & 1 & 0 \\ 1 & 2 & 1 & 3 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

and the last row shows that the system is inconsistent.

(c) Now write a matrix equation that can be used to find the least-squares solution to the equation you wrote in (a). Fully simplify any matrix products that occur in your equation, but *do not (yet) attempt to solve the equation.*

Solution. We can use the normal equation $A^\top A\vec{x} = A^\top \vec{b}$ to get the least-squares solution to $A\vec{x} = \vec{b}$ found above.

$$A^\top A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 3 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 14 & 5 \\ 14 & 30 & 10 \\ 5 & 10 & 4 \end{bmatrix}$$

$$A^\top \vec{b} = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 3 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 17 \\ 36 \\ 11 \end{bmatrix}$$

so the matrix equation is:

$$\begin{bmatrix} 7 & 14 & 5 \\ 14 & 30 & 10 \\ 5 & 10 & 4 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 17 \\ 36 \\ 11 \end{bmatrix}$$

(d) Now, solve your equation using methods taught in this course. (You can use a matrix calculator to check your answer, but you must be able to solve this problem by hand.)

Solution.
Performing row operations we find the reduced row echelon form of our augmented coefficient matrix:

$$\begin{bmatrix} 7 & 14 & 5 & \mid & 17 \\ 14 & 30 & 10 & \mid & 36 \\ 5 & 10 & 4 & \mid & 11 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & \mid & -7/3 \\ 0 & 1 & 0 & \mid & 1 \\ 0 & 0 & 1 & \mid & -8/3 \end{bmatrix}$$

We conclude that the least-squares solution to the problem is $A = -7/3, B = 1, C = -8/3$, so that the equation of the "plane of best fit" is

$$y = -\frac{7}{3}x + y - \frac{8}{3}.$$

(e) **(Recreational:)**[1] Use 3-D graphing software such as GeoGebra or Desmos 3D to plot the four points and graph the "plane of best fit" through them.

Problem 2.

(a) Which of the following is an inner product in $\mathcal{P}_2$? Explain.

(i) $\langle f, g\rangle = f(1)g(2) + f(2)g(1) + f(3)g(3)$

Solution. This is not an inner product as it does not satisfy positive definiteness. For any quadratic polynomial $f(x)$, we would have

$$\langle f, f\rangle = f(1)f(2) + f(2)f(1) + f(3)f(3) = 2f(1)f(2) + f(3)^2$$

. This can be zero even for a nonzero polynomial; for example, the quadratic polynomial $f(x) = (x-1)(x-3)$ is a nonzero polynomial for which $\langle f, f\rangle = 0$, which shows that this does not define an inner product.

(ii) $\langle f, g\rangle = f(1)g(1) + f(2)g(2) + f(3)g(3)$

Solution. This is an inner product.
Symmetry: $\langle f, g\rangle = f(1)g(1) + f(2)g(2) + f(3)g(3) = \langle g, f\rangle$ by commutativity of multiplication
Linearity:

$$\begin{aligned} \langle f + h, g\rangle &= (f(1) + h(1))g(1) + (f(2) + h(2))g(2) + (f(3) + h(3))g(3) \\ &= f(1)g(1) + f(2)g(2) + f(3)g(3) + h(1)g(1) + h(2)g(2) + h(3)g(3) \\ &= \langle f, g\rangle + \langle h, g\rangle \end{aligned}$$

$$\langle cf, g\rangle = cf(1)g(1) + cf(2)g(2) + cf(3)g(3) = c\langle f, g\rangle$$

[1] Recreational problems are for your own interest only; you are not required to submit your solutions, and if you do, they will not be graded. However, you may find that doing these problems helps you gain a better understanding of what we are doing in this somewhat abstract unit.

Positive definiteness: Note that

$$\langle f, f\rangle = f(1)^2 + f(2)^2 + f(3)^2$$

which is a sum of non-negative terms. This implies that $\langle f, f\rangle \geq 0$ for all f, and $\langle f, f\rangle = 0$ if and only if $f(1), f(2)$ *and* $f(3)$ are all 0. Since f is a polynomial of degree at most 2, it can have at most 2 roots unless it is the constant zero function. Thus this inner product space *is* positive definite.

(b) Let $V = C^\infty[-1, 1]$, the vector space of smooth functions on the interval $[-1, 1]$. Which of the following is an inner product in V? Explain.

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

Solution. This is not an inner product. Although it is symmetric and linear in both arguments, it is not positive definite. Consider the function $f(x) = x$. Then $\langle f, f\rangle = \int_{-1}^{1} x^3\, dx = 0$.

(ii) $\langle f, g\rangle = \displaystyle\int_{-1}^{1} x^2 f(x)g(x)\, dx$

Solution. This is an inner product.
Symmetry: $\langle f, g\rangle = \int_{-1}^{1} x^2 f(x)g(x)\, dx = \int_{-1}^{1} x^2 g(x)f(x)\, dx = \langle g, f\rangle$.
Linearity: $\langle f + h, g\rangle = \int_{-1}^{1} x^2(f(x) + h(x))g(x)\, dx = \int_{-1}^{1} x^2 f(x)g(x)\, dx + \int_{-1}^{1} x^2 h(x)g(x)\, dx = \langle f, g\rangle + \langle h, g\rangle$.
Positive definiteness: Observe that for any function $f(x)$, it is always the case that $x^2 f(x)^2$ is non-negative. So $\langle f, f\rangle = \int_{-1}^{1} x^2 f(x)^2\, dx$ is non-negative. Now the integral of a non-negative can be zero if and only if the function itself is identically zero over the entire interval of integration, so $x^2 f(x)^2$ must be zero on all of $[-1, 1]$. Therefore $f(x) = 0$ on all of $[-1, 1]$.

Problem 3. Let $V = C^\infty\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, the vector space of smooth functions on the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, and consider the inner product defined by $\langle f, g\rangle = \displaystyle\int_{-\pi/2}^{\pi/2} f(x)g(x)\sin^2(x)\, dx$. (You do not need to show that this is an inner product, but make sure that you would be able to do so if it were an exam question!) Let $W = \operatorname{span}(1, x, x^2)$.

In what follows, you may feel free to use an online integral calculator (e.g. Wolfram Alpha) to evaluate any difficult integrals[2], but make sure that your work shows clearly *what integrals* you are computing, and how you are making use of the results. Results may be expressed using either exact expressions (e.g., $\pi/\sqrt{2}$) or decimal approximations (e.g., 2.2214), but if you use decimal approximations, please retain at least four digits' worth of precision.

(a) Compute each of the following.

(i) $\langle 1, x\rangle$
(ii) $\|1\|$
(iii) $\|x\|$

[2]Despite what it may seem sometimes, the goal of the Inner Products unit is not to test your integration skills.

Solution.
(i) $\langle 1, x\rangle = \int_{-\pi/2}^{\pi/2} x\sin^2 x\, dx = 0$ (the easiest way to compute this is to notice that it is the integral of an odd function over a symmetric interval).
(ii) $\|1\|^2 = \langle 1, 1\rangle = \int_{-\pi/2}^{\pi/2} \sin^2(x)\, dx = \pi/2$, so $\|1\| = \sqrt{\pi/2} \approx 1.2533$.
(iii) $\|x\|^2 = \langle x, x\rangle = \int_{-\pi/2}^{\pi/2} x^2\sin^2(x)\, dx = \frac{\pi}{24}(6+\pi^2)$, so $\|x\| = \sqrt{\frac{\pi}{24}(6+\pi^2)} \approx 1.4413$.

(b) Find a basis $\mathcal{U}$ for the subspace W that is orthonormal relative to the given inner product.

Solution. We use Gramm-Schmidt on the initial basis $(1, x, x^2)$ to find an orthonormal basis of functions $\mathcal{U} = (g_1, g_2, g_3)$.

First, we take $g_1(x) = \frac{1}{\|1\|} = \sqrt{\frac{2}{\pi}} \approx= 0.7979$.

Second, we note from part (a)(i) that 1 is orthogonal to x, and hence g_1 is orthogonal to x as well. So we can take for our second unit vector $g_2(x) = \frac{x}{\|x\|} = \sqrt{\frac{24}{\pi(6+\pi^2)}}\, x \approx 0.6938x$.

For our third unit vector, we calculate

$$\begin{aligned}
&x^2 - \left\langle x^2, \sqrt{\frac{2}{\pi}}\right\rangle \sqrt{\frac{2}{\pi}} - \left\langle x^2, \sqrt{\frac{24}{\pi(6+\pi^2)}}\, x\right\rangle \sqrt{\frac{24}{\pi(6+\pi^2)}}\, x \\
&= x^2 - \frac{2}{\pi}\left\langle x^2, 1\right\rangle - \frac{24}{\pi(6+\pi^2)}\left\langle x^2, x\right\rangle x^2 \\
&= x^2 - \frac{2}{\pi}\frac{\pi}{24}(6+\pi^2) - 0 \\
&= x^2 - \frac{6+\pi^2}{12} \\
&\approx x^2 - 1.3224
\end{aligned}$$

which we need to then divide by its own magnitude:

$$\|x^2 - 1.3224\| = \left\langle x^2 - 1.3224, x^2 - 1.3224\right\rangle = \int_{-\pi/2}^{\pi/2} (x^2 - 1.3224)^2 \sin^2(x)\, dx \approx 0.6850$$

so we take

$$g_3(x) = \frac{1}{\sqrt{0.6850}}(x^2 - 1.3224) \approx 1.208x^2 - 1.5978$$

(c) Let $h \in C^\infty\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ be the function defined by $h(x) = e^x$ for all $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Compute $\operatorname{proj}_W h$.

Solution. We just have to hold our nose and compute:

$$\langle e^x, 0.7979\rangle\, 0.7979 + \langle e^x, 0.6938x\rangle\, 0.6938x + \left\langle e^x, 1.208x^2 - 1.5978\right\rangle \left(1.208x^2 - 1.5978\right)$$

The first inner product is:

$$\langle e^x, 0.7979 \rangle = \int_{-\pi/2}^{\pi/2} 0.7979 e^x \sin^2 x \, dx \approx 2.20345,$$

so $\langle e^x, 0.7979 \rangle \, 0.7979 \approx (2.20345)(0.7979) = 1.7581$.

The second inner product is:

$$\langle e^x, 0.6938x \rangle = \int_{-\pi/2}^{\pi/2} 0.6938 x e^x \sin^2 x \, dx \approx 1.87641,$$

so $\langle e^x, 0.6938x \rangle \, 0.6938x \approx (1.87641)(0.6938x) = 1.30185x$.

The third inner product is:

$$\left\langle e^x, 1.208x^2 - 1.5978 \right\rangle = \int_{-\pi/2}^{\pi/2} \left(1.208x^2 - 1.5978\right) e^x \sin^2 x \, dx \approx 0.50965$$

so $\left\langle e^x, 1.208x^2 - 1.5978 \right\rangle \left(1.208x^2 - 1.5978\right) = (0.50965)\left(1.208x^2 - 1.5978\right) = 0.6167x^2 - 0.8143$.

Finally, the projection of e^x onto our subspace W is

$$1.7581 + 1.30185x + 0.6167x^2 - 0.8143 = 0.9438 + 1.30185x + 0.6167x^2$$

(d) **(Recreational:)** Repeat parts (a)–(c), this time using the simpler inner product $\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, dx$.

(e) **(Recreational:)** Use graphing software (e.g., Desmos) to plot the function h and the two different projections you found in (c) and (d) over the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, all on the same axes. How do these three functions compare? Which of the two projections does a "better job" of approximating h (and in what sense is it "better"?) What are some situations in which you might choose to use one inner product rather than the other?

Solution. The approximation using the inner product from (d) does a very good job approximating $h(x) = e^x$ on the interval $[-1, 1]$, but outside of that interval the approximation is not very good. This is because the projection minimizes the value of

$$\left\langle h(x) - \operatorname{proj}_W h, h(x) - \operatorname{proj}_W h \right\rangle,$$

and this inner product only "looks at" values of the function in the interval $[-1, 1]$.

In contrast, the approximation using the inner product from (a)-(c) uses the entire interval $[-\pi/2, \pi/2]$. Moreover, the function $\sin^2 x$ is close to 1 for values of x on the outer edges of that interval, and is close to zero for values of x within the inner core of the interval. As a result, the inner product "weighs" gaps on the outer edges more heavily than gaps on the inside. The result of this is that using the more complicated inner product produces an approximation that minimizes gaps near the outer edge of the interval, but does not try to minimize gaps near the center.

The take-away from this is that when you are trying to approximate a function, you should use an inner product that measures the thing you are trying to make small!