Homework 5 Partial Solutions

Homework 5 Problems

5.1

1.

(a)

$\cos(\theta) = \frac{\boldsymbol{w}^T\boldsymbol{v}}{||\boldsymbol{w}||||\boldsymbol{v}||} = 1$ so $\boldsymbol{w}$ and $\boldsymbol{v}$ are in the same direction.

This is clear since $3(2,1,3) = (6,3,9)$.

5. $y = 2x$ is the same as $U = \text{span}\{\boldsymbol{u}\}$, where $\boldsymbol{u} = (1,2)$. The point in U closest to $\boldsymbol{v} = (5,2)$ is the projection $P_U\boldsymbol{v}$ where P_U is the projection map onto U, this has matrix $P_U = A(A^TA)^{-1}A^T$, where $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $U = \text{CS}(A)$.

Notice, $A(A^TA)^{-1}A^T\boldsymbol{v} = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{u})^{-1}(\boldsymbol{u}^T\boldsymbol{v}) = \frac{\langle \boldsymbol{v},\boldsymbol{u}\rangle}{\langle \boldsymbol{u},\boldsymbol{u}\rangle}\boldsymbol{u} = \text{proj}_{\boldsymbol{u}}(\boldsymbol{v})$, which is how the text defines the projection of $\boldsymbol{v}$ onto $\boldsymbol{u}$.

$$P_U = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

so $P_U\boldsymbol{v} = \frac{1}{5}(9,18)$ or if you like $\text{proj}_{\boldsymbol{u}}(\boldsymbol{v}) = \frac{9}{5}(1,2)$.

13. Let $\boldsymbol{v}$ and $\boldsymbol{u}$ be vectors in any inner product space. We have

$$\begin{aligned}
\|\boldsymbol{v}+\boldsymbol{u}\|^2 &= \langle \boldsymbol{v}+\boldsymbol{u}, \boldsymbol{v}+\boldsymbol{u}\rangle \\
&= \langle \boldsymbol{v}, \boldsymbol{v}+\boldsymbol{u}\rangle + \langle \boldsymbol{u}, \boldsymbol{v}+\boldsymbol{u}\rangle \\
&= \langle \boldsymbol{v}, \boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{u}\rangle + \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{u}, \boldsymbol{u}\rangle \\
&= \|\boldsymbol{v}\|^2 + \|\boldsymbol{u}\|^2 + 2\langle \boldsymbol{v}, \boldsymbol{u}\rangle \\
&\le \|\boldsymbol{v}\|^2 + \|\boldsymbol{u}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{u}\| && \text{(Cauchy's Theorem)} \\
&= (\|\boldsymbol{v}\| + \|\boldsymbol{u}\|)^2
\end{aligned}$$

So $\|\boldsymbol{v}+\boldsymbol{u}\| \le \|\boldsymbol{v}\| + \|\boldsymbol{u}\|$.

Equality will hold when $\langle \boldsymbol{v}, \boldsymbol{u}\rangle = 0$, this would mean

$$(\langle \boldsymbol{u}, \boldsymbol{v}\rangle)^2 = \langle \boldsymbol{v}, \boldsymbol{v}\rangle\langle \boldsymbol{u}, \boldsymbol{u}\rangle$$

Assuming $\|\boldsymbol{u}\| \neq 0 \neq \|\boldsymbol{v}\|$, then we have

$$\begin{aligned}
\langle \|\boldsymbol{v}\|\boldsymbol{u} - \|\boldsymbol{u}\|\boldsymbol{v}, \|\boldsymbol{v}\|\boldsymbol{u} - \|\boldsymbol{u}\|\boldsymbol{v}\rangle &= \|\boldsymbol{v}\|\langle \boldsymbol{u}, \|\boldsymbol{v}\|\boldsymbol{u} - \|\boldsymbol{u}\|\boldsymbol{v}\rangle - \|\boldsymbol{u}\|\langle \boldsymbol{v}, \|\boldsymbol{v}\|\boldsymbol{u} - \|\boldsymbol{u}\|\boldsymbol{v}\rangle \\
&= \|\boldsymbol{v}\|^2\langle \boldsymbol{u}, \boldsymbol{u}\rangle - \|\boldsymbol{v}\|\|\boldsymbol{u}\|\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \|\boldsymbol{u}\|\|\boldsymbol{v}\|\langle \boldsymbol{v}, \boldsymbol{u}\rangle + \|\boldsymbol{u}\|^2\langle \boldsymbol{v}, \boldsymbol{v}\rangle \\
&= 2\|\boldsymbol{v}\|^2\|\boldsymbol{v}\|^2 - 2(\|\boldsymbol{u}\|\|\boldsymbol{v}\|)^2 = 0
\end{aligned}$$

So $\|\boldsymbol{v}\|\boldsymbol{u} - \|\boldsymbol{u}\|\boldsymbol{v} = \mathbf{0}$, but this means that $\boldsymbol{u}$ and $\boldsymbol{v}$ differ by a scalar multiple.

18.

(a) Show that

$$p_{\boldsymbol{y}}(\boldsymbol{x}) = \frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\langle \boldsymbol{y}, \boldsymbol{y}\rangle}\boldsymbol{y}$$

is the orthogonal projection of $\boldsymbol{x}$ onto $\boldsymbol{y}$. That is, $\boldsymbol{x} - p_{\boldsymbol{y}}(\boldsymbol{x}) \perp \boldsymbol{y}$.

Notice the usual inner product is $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{y}^T\boldsymbol{x}$, so this does solve the problem.

Clearly

$$\langle \boldsymbol{x} - p_{\boldsymbol{y}}(\boldsymbol{x}), \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle - \frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\langle \boldsymbol{y}, \boldsymbol{y}\rangle}\langle \boldsymbol{y}, \boldsymbol{y}\rangle = 0$$

(b) In arguing for (13) above we see $\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 \iff \langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$, so here we have

$$\|\boldsymbol{x}\|^2 = \|\boldsymbol{x} - p_{\boldsymbol{y}}(\boldsymbol{x}) + p_{\boldsymbol{y}}(\boldsymbol{x})\|^2 = \|\boldsymbol{x} - p_{\boldsymbol{y}}(\boldsymbol{x})\|^2 + \|p_{\boldsymbol{y}}(\boldsymbol{x})\|^2 = 8^2 + 6^2 = 10^2$$

So $\|\boldsymbol{x}\| = 10$.

5.2

1.

(d) Here

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix} \quad \text{and} \quad \operatorname{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

So we read off a basis for the four subspaces:

$\mathrm{RS}(A) = \mathrm{CS}(A^T)$: A basis is $\{(1,0,0,0), (0,1,0,0), (0,0,1,1)\}$ (the non-zero rows of $\operatorname{rref}(A)$)

$\mathrm{CS}(A)$: A basis is $\{(1,0,0,1), (0,1,0,1), (0,1,1,2)\}$ (the pivot columns of A).

$\mathrm{NS}(A)$: A basis is $\{(0,0,0,1)\}$ since we see that the solutions to $A\boldsymbol{x} = \mathbf{0}$ are of the form $\boldsymbol{x} = t\,(0,0,-1,1)$.

Note: $\mathrm{NS}(A) \perp \mathrm{CS}(A^T)$ as we know must happen.

5. If A is 3×2 and of rank 2, then $\mathrm{RS}(A)$ is a 2-dimensional subspace of $\mathbb{R}^2$, hence $\mathrm{RS}(A) = \mathbb{R}^2$. $\mathbb{R}^3 = \mathrm{NS}(A^T) \oplus \mathrm{RS}(A^T)$ and $\mathrm{RS}(A^T) = \mathrm{CS}(A)$ is a 2-dimensional subset of $\mathbb{R}^3$, a plane. So $\mathrm{RS}(A^T)$ is a plane in $\mathbb{R}^3$ and $\mathrm{NS}(A^T)$ is the line normal to that plane.

13.

(a) Let $\boldsymbol{x} \in \mathrm{NS}(A^TA)$. $Ax \in \mathrm{rng}(A)$ by definition of $\mathrm{rng}(A)$, no assumption necessary. Of course $\boldsymbol{x} \in \mathrm{NS}(A^TA)$ means $(A^TA)\boldsymbol{x} = \mathbf{0}$, but $(A^TA)\boldsymbol{x} = A^T(A\boldsymbol{x}) = \mathbf{0} \implies A\boldsymbol{x} \in \mathrm{NS}(A^T)$.

(b) For any $\boldsymbol{x}$, $\boldsymbol{x} \in \mathrm{NS}(A) \implies A\boldsymbol{x} = \mathbf{0} \implies A^TA\boldsymbol{x} = 0$. So

$$\mathrm{NS}(A) \subseteq \mathrm{NS}(A^TA)$$

So let $\boldsymbol{x} \in \mathrm{NS}(A^TA)$, then $A\boldsymbol{x} \in \mathrm{NS}(A^T) \cap \mathrm{rng}(A)$ by (a). But $\mathrm{rng}(A) = \mathrm{CS}(A) = \mathrm{RS}(A^T)$ and so $A\boldsymbol{x} \in \mathrm{NS}(A^T) \cap \mathrm{RS}(A^T)$. But $\mathrm{NS}(A^T) \perp \mathrm{RS}(A^T)$ so $\mathrm{NS}(A^T) \cap \mathrm{RS}(A^T) = \{\mathbf{0}\}$. Thus $A\boldsymbol{x} = \mathbf{0}$ and $\boldsymbol{x} \in \mathrm{NS}(A)$. So we have

$$\mathrm{NS}(A^TA) \subseteq \mathrm{NS}(A)$$

So we have $\mathrm{NS}(A^TA) = \mathrm{NS}(A)$ as desired.

(c) If A is $m \times n$, then A^TA is $n \times n$ and the rank-nullity theorem gives

$$\mathrm{rank}(A) + \dim(\mathrm{NS}(A)) = n = \mathrm{rank}(A^TA) + \dim(\mathrm{NS}(A^TA))$$

By (b), $\dim(\mathrm{NS}(A)) = \dim(\mathrm{NS}(A^TA))$ so $\mathrm{rank}(A) = \mathrm{rank}(A^TA)$.

(d) If A has independent columns, then $\mathrm{rank}(A) = n = \mathrm{rank}(A^TA)$ so A^TA has full rank and is invertible.

14 A is $m \times n$, B is $n \times r$. and $C = AB$.

(a) $\boldsymbol{x} \in \mathrm{NS}(B) \implies B\boldsymbol{x} = \mathbf{0} \implies AB\boldsymbol{x} = \mathbf{0} \implies \boldsymbol{x} \in \mathrm{NS}(C)$, so $\mathrm{NS}(B) \subseteq \mathrm{NS}(C)$.

(b) $\mathrm{NS}(C) \oplus \mathrm{NS}(C)^\perp = \mathbb{R}^r = \mathrm{NS}(B) \oplus \mathrm{NS}(B)^\perp$, so $\mathrm{rng}(C^T) = \mathrm{CS}(C^T) = \mathrm{NS}(C)^\perp \subseteq \mathrm{NS}(B)^\perp = \mathrm{CS}(B^T) = \mathrm{rng}(B^T)$.

15. By definition $W = U \oplus V$ iff $W = U + V$ for any $\boldsymbol{w} \in W$, there are unique $\boldsymbol{u}$ and $\boldsymbol{v}$ so that $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$.

If $\boldsymbol{z} \in U \cap V$ and $\boldsymbol{z} \neq \mathbf{0}$, then $\boldsymbol{z} = \boldsymbol{u} + \boldsymbol{v}$, where there are two options $\boldsymbol{u} = \boldsymbol{z}$ and $\boldsymbol{v} = \mathbf{0}$ or $\boldsymbol{u} = \mathbf{0}$ and $\boldsymbol{v} = \boldsymbol{z}$.

16. Let $\mathrm{rank}(A) = k$ and $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ be a basis for $\mathrm{RS}(A) = \mathrm{rng}(A^T)$. Then to show that $\{A\boldsymbol{x}_1, \ldots, A\boldsymbol{x}_k\}$ is a basis for $\mathrm{rng}(A) = \mathrm{CS}(A)$ we only need to check independence, since we know $\dim(\mathrm{RS}(A)) = \dim(\mathrm{CS}(A)) = \mathrm{rank}(A)$. Suppose $\sum_{i=1}^k \alpha_i A\boldsymbol{x}_i = \mathbf{0}$, then $A\left(\sum_{i=1}^k \alpha_i \boldsymbol{x}_i\right) = \mathbf{0}$, so $\sum_{i=1}^k \alpha_i \boldsymbol{x}_i \in \mathrm{NS}(A)$. But $\mathrm{NS}(A) \cap \mathrm{rng}(A^T) = \{\mathbf{0}\}$ and thus $\sum_{i=1}^k \alpha_i \boldsymbol{x}_i = \mathbf{0}$ and so $\alpha_i = 0$ for all i.

5.3

3.

(b) Find first the orthogonal projection of $\boldsymbol{b}$ onto $\mathrm{CS}(A)$. The first two columns of A are a basis for $\mathrm{CS}(A)$ so we can use those two to find the projection $\hat{\boldsymbol{b}} = B(B^TB)^{-1}B^T$ where

$B = A(:, 1:2)$. We find $\hat{\boldsymbol{b}} = (3, 1, 4)^T$ now we can ask what $\hat{\boldsymbol{x}}$ satisfy $A\hat{\boldsymbol{x}} = (3, 1, 4)^T$. Solving this you find solutions have the form

$$\hat{\boldsymbol{x}} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$

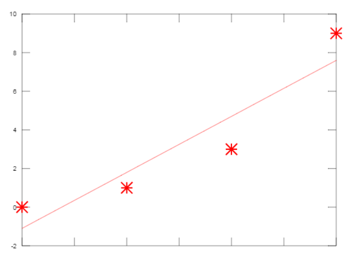
5. Trying to find $(\alpha, \beta)^T$ so that

$$\underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}_{\hat{\boldsymbol{x}}} = \underbrace{\begin{bmatrix} 0 \\ 1 \\ 3 \\ 9 \end{bmatrix}}_{b}$$

For this $\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (A^T A)^{-1} A^T \boldsymbol{b}$ and we get:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1.8 \\ 2.9 \end{bmatrix}$$

So the equation is $y = 2.9x + 1.8$.



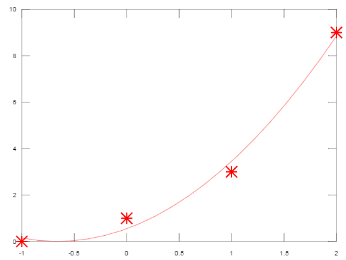
6. Trying to find $(\alpha, \beta)^T$ so that

$$\underbrace{\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}}_{\hat{\boldsymbol{x}}} = \underbrace{\begin{bmatrix} 0 \\ 1 \\ 3 \\ 9 \end{bmatrix}}_{b}$$

For this $\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (A^T A)^{-1} A^T \boldsymbol{b}$ and we get:

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0.55 \\ 1.65 \\ 1.25 \end{bmatrix}$$

So the equation is $y = 1.25x^2 + 1.65x + 0.55$.



13. We know that if A is $m \times n$ and $\boldsymbol{b} \in \mathbb{R}^m$, then there is a unique $\hat{\boldsymbol{b}} \in \text{rng}(A) = \text{CS}(A)$ so that $\hat{\boldsymbol{b}} - \boldsymbol{b} \perp \text{rng}(A)$ and hence $\|\hat{\boldsymbol{b}} - \boldsymbol{b}\|$ is minimal among all $\|\boldsymbol{y} - \boldsymbol{b}\|$ for all $\boldsymbol{y} \in \text{rng}(A)$.

You should understand why $\hat{\boldsymbol{b}} = A(A^TA)^{-1}A^Tb$. So that $\hat{x} = (A^TA)^{-1}A^Tb$ is such that $\|A\hat{\boldsymbol{x}} - \boldsymbol{b}\|$ is minimized and what you know is $A^TA\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$.

We are assuming that $\hat{\boldsymbol{x}}$ is a least-squares solution to $A\boldsymbol{x} = \boldsymbol{b}$, so $A^TA\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$. There are two things to prove:

(1) If $\boldsymbol{y}$ is is a least-squares solution to $A\boldsymbol{x} = \boldsymbol{b}$, then $\boldsymbol{y} = \hat{\boldsymbol{x}} + \boldsymbol{z}$ for $\boldsymbol{z} \in \text{NS}(S)$.

We know $A^TA\boldsymbol{y} = A^T\boldsymbol{b} = A^TA\hat{\boldsymbol{x}}$ so $A^TA(\boldsymbol{y} - \hat{\boldsymbol{x}}) = 0$. Let $\boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{x}}$, then $\boldsymbol{z} \in \text{NS}(A^TA) = \text{NS}(A)$ and $\boldsymbol{y} = \hat{\boldsymbol{x}} + \boldsymbol{z}$ as desired.

(2) If $\boldsymbol{y} = \hat{\boldsymbol{x}} + \boldsymbol{z}$ for some $\boldsymbol{z} \in \text{NS}(A)$, then $\boldsymbol{y}$ is a least square-solution to $A\boldsymbol{x} = \boldsymbol{b}$.

$A^TA(\hat{\boldsymbol{x}} + \boldsymbol{z}) = A^TA\hat{\boldsymbol{x}} + A^TA\boldsymbol{z} = \boldsymbol{b} + \boldsymbol{0} = \boldsymbol{b}$. So $\hat{\boldsymbol{x}} + \boldsymbol{z}$ is a least-squares solution to $A\boldsymbol{x} = \boldsymbol{b}$.

14. We want to fit a circle to the data $(-1, -2), (0, 2.4), (1.1, -4), (2.4, -1.6)$. An arbitrary circle would satisfy

$$x^2 + y^2 + ax + by + c = 0$$

or

$$ax + by + c = -(x^2 + y^2)$$

So we are trying to solve $A\boldsymbol{x} = \boldsymbol{b}$ where

$$\underbrace{\begin{bmatrix} -1 & -2 & 1 \\ 0 & 2.4 & 1 \\ 1.1 & -4 & 1 \\ 2.4 & -1.6 & 1 \end{bmatrix}}_{A} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \underbrace{\begin{bmatrix} -((-1)^2 + (-2)^2) \\ -((0)^2 + (2.4)^2) \\ -((1.1)^2 + (-4)^2) \\ -((2.4)^2 + (1.6)^2) \end{bmatrix}}_{\mathbf{b}}$$

Here is MATLAB code for this and here it is done with Desmos.

5.4

4. Given

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

(a)

$$\begin{aligned}\langle A, B\rangle = (1)(-4)+&(2)(1)+(2)(1)+(1)(-3)+(0)(3)\\ &+(2)(2)+(3)(1)+(1)(-2)+(1)(-2)=0\end{aligned}$$

So A and B are orthogonal.

(b)

$$\|A\|_F^2 = (1)^2+(2)^2+(2)^2+(1)^2+(0)^2+(2)^2+(3)^2+(1)^2+(1)^2 = 25$$

(c)

$$\|A\|_F^2 = (-4)^2+(1)^2+(1)^2+(-3)^2+(3)^2+(2)^2+(1)^2+(-2)^2+(-2)^2 = 49$$

(d) $\|A+B\|_F^2 = \|A\|_F^2 + \|B\|_F^2 = 25+49 = 74$, since $A \perp_F B$.

8.

(a)

$$\cos(\theta) = \frac{\langle 1, x\rangle}{\|1\|\cdot\|x\|} = \frac{\int_0^1 (1\cdot x)\,dx}{\left(\int_0^1 1^2\,dx\right)^{1/2}\cdot\left(\int_0^1 x^2\,dx\right)^{1/2}} = \frac{1/2}{(1)^{1/2}\cdot(1/3)^{1/2}} = \sqrt{3}/2$$

So $\theta = \pi/6$. Of course, this really doesn't mean anything, the relevant thing here is $\frac{\langle 1,x\rangle}{\|1\|\cdot\|x\|}$.

(b) $p = p_x(1) = \frac{\langle 1,x\rangle}{\langle x,x\rangle}\,x = \frac{\int_0^1 x\,dx}{\int_0^1 x^2\,dx} = \frac{1/2}{1/3}x = (3/2)x$.

Check that $1-p \perp p$, $\langle p, 1-p\rangle = \int_0^1 (3/2)x(1-(3/2)x)\,dx = \int_0^1 (3/2)x - (3/2)^2x^2\,dx = (3/2)(1/2) - (1/3)(3/2)^2 = 0$!

(c)

$$\begin{aligned}\|1\|^2 &= \langle 1, 1\rangle = \int_0^1 1\,dx = 1\\ \|p\|^2 &= \|(3/2)x\|^2 = (3/2)^2\langle x, x\rangle = \int_0^1 x^2\,dx = (3/2)^2\frac{1}{3} = 3/4\\ \|1-p\|^2 &= \int_0^1 (1-(3/2)x)^2\,dx = (2/3)\int_1^{-1/2} u^2\,du\\ &= (2/3)(1/3)((-1/2)^3 - (1^3)) = (2/3)(1/3)(-9/8) = 1/4\end{aligned}$$

Now

$$\|p\|^2 + \|1-p\|^2 = 3/4 + 1/4 = 1 = \|1\|^2 = \|p + (1-p)\|^2$$

Verifying Pythagoras.

10. Let $x_i = (i-3)/2$ for $i = 1,\ldots,5$. Define $\langle p, q\rangle = \sum_{i=1}^5 p(x_i)q(x_i)$ for $p, q \in \mathbb{P}_5[x]$.

$$\langle x, x^2\rangle = \sum_{i=1}^5 (i-3)/2\cdot(i-3)/2^2 = \frac{1}{8}\sum_{i=1}^5 (i-3)^3 = \frac{1}{8}((-2)^3+(-1)^3+(0)^3+(1)^3+(2)^3) = 0$$

So $x \perp x^2$ as desired.

5.5

29 Use $\langle f,g\rangle = \int_{-1}^{1} f\cdot g\,dx$

(a) $\langle 1,x\rangle = \int_{-1}^{1} 1\cdot x\,dx = \frac{1}{2}x^2\Big|_{-1}^{1} = 0$. So $1\perp x$.

(b)

$$\|1\|^2 = \langle 1,1\rangle = \int_{-1}^{1} 1\cdot 1\,dx = x\Big|_{-1}^{1} = 2$$
$$\|x\|^2 = \langle x,x\rangle = \int_{-1}^{1} x^2\,dx = \frac{1}{3}x^3\Big|_{-1}^{1} = \frac{2}{3}$$

(c) To find the best approximation to $x^{1/3}$ we project $x^{1/3}$ onto span$\{1,x\}$

$$p_1(x^{1/3}) = \frac{\langle x^{1/3},1\rangle}{\langle 1,1\rangle}1 = \frac{\int_{-1}^{1} x^{1/3}\,dx}{2} = \frac{(3/4)x^{4/3}\Big|_{-1}^{1}}{2} = 0$$
$$p_x(x^{1/3}) = \frac{\langle x^{1/3},x\rangle}{\langle x,x\rangle}x = \frac{\int_{-1}^{1} x^{4/3}\,dx}{(2/3)}x = \frac{(3/7)x^{7/3}\Big|_{-1}^{1}}{(2/3)}x = 2(3/7)(3/2)x = 9/7\,x$$

So the projection of $x^{1/3}$ onto span$\{1,x\}$ is $0\cdot 1 + 9/7\cdot x = 9/7\,x$

Here is a Desmos graph to illustrate this.

30

(a) Let $f_1 = 1$ and $f_2 = 2x-1$, $\langle f_1,f_2\rangle = \int_0^1 f_1\cdot f_2\,dx = \int_0^1 (2x-1)\,dx = x^2 - x\Big|_0^1 = 0$, so $f_1\perp f_2$.

(b) $\|f_1\|_2^2 = \langle f_1,f_1\rangle = \int_0^1 f_1^2\,dx = \int_0^1 (1)^2\,dx = x\Big|_0^1 = 1$ and $\|f_2\|_2^2 = \int_0^1 (2x-1)^2\,dx = \frac{1}{2}\int_{-1}^{1} u^2\,du = \frac{1}{2}\frac{u^3}{3}\Big|_{-1}^{1} = \frac{1}{3}$. So $\|f_1\| = 1$ and $\|f_2\| = \frac{1}{\sqrt{3}}$, so the unit vector in the direction of f_2 is $\sqrt{3}f_2$.

(c) The projection of $g(x) = \sqrt{x}$ onto span$\{f_1,f_2\}$ is $\langle g,f_1\rangle\cdot f_1 + \langle g,\sqrt{3}f_2\rangle\cdot\sqrt{3}f_2 = \langle g,f_1\rangle\cdot f_1 + \langle g,f_2\rangle\cdot 3f_2$. We have

$$\langle g,f_1\rangle = \int_0^1 x^{1/2}\,dx = \frac{2}{3}x^{3/2}\Big|_0^1 = \frac{2}{3}$$
$$\begin{aligned}\langle g,f_2\rangle &= \int_0^1 x^{1/2}(2x-1)\,dx\\ &= \int_0^1 2x^{3/2} - x^{1/2}\,dx\\ &= \left(2\cdot\frac{2}{5}x^{5/2} - \frac{2}{3}x^{3/2}\right)\Big|_0^1 = \frac{2}{15}\end{aligned}$$

So the projection of $g(x) = \sqrt{x}$ onto span$\{f_1,f_2\}$ is

$$\hat{g} = \frac{2}{3}\cdot f_1 + \frac{2}{15}\cdot 3\cdot f_2 = \frac{2}{3} + \frac{2}{5}(2x-1) = \frac{4}{5}x + \frac{4}{15}$$

See demo here

1 5.6

4. The strategy here is simple:

- Start with $\mathcal{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\} = \{1, x, x^2\}$.
- $\boldsymbol{u}_1 = v_1$
- $\boldsymbol{q}_1 = \boldsymbol{u}_1/||u_1||$
- $\boldsymbol{u}_2 = \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{q}_1 \rangle \, \boldsymbol{q}_1$
- $\boldsymbol{q}_2 = \boldsymbol{u}_2/||\boldsymbol{u}_2||$
- $\boldsymbol{u}_3 = \boldsymbol{v}_3 - \langle \boldsymbol{v}_3, \boldsymbol{q}_1 \rangle \, \boldsymbol{q}_1 - \langle \boldsymbol{v}_3, \boldsymbol{q}_2 \rangle \, \boldsymbol{q}_2$
- $\boldsymbol{q}_3 = \boldsymbol{u}_3/||\boldsymbol{u}_3||$

That is the idea, here are the details. This is how all simple Gram-Schmidt proceeds.

Choose $\boldsymbol{u}_1 = 1$, then $\langle 1, 1 \rangle = \int_0^1 1 \, dx = 1$ so this is already normalized and so set $\boldsymbol{q}_1 = \boldsymbol{u}_1$.

Set $\boldsymbol{u}_2 = x - \langle x, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 = x - \int_0^1 x \, dx = x - \frac{1}{2}$. Now $||\boldsymbol{u}_2||^2 = \int_0^1 \left(x - \frac{1}{2}\right)^2 \, dx = \int_0^1 x^2 - x + \frac{1}{4} \, dx = \frac{1}{3}x^3 - \frac{1}{2}x^2 + \frac{1}{4}x\Big|_{x=0}^{x=1} = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}$. So $\boldsymbol{q}_2 = \sqrt{12}\left(x - \frac{1}{2}\right) = \sqrt{3}(2x-1)$.

Finally, $\boldsymbol{u}_3 = x^2 - \langle x^2, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2 - \langle x^2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1$. We have $\langle x^2, \boldsymbol{q}_2 \rangle = \int_0^1 \sqrt{3}(2x-1) \, x^2 \, dx = \sqrt{3}\left(\frac{1}{2}x^4 - \frac{1}{3}x^3\right)\Big|_{x=0}^{x=1} = \frac{1}{2\sqrt{3}}$. So $\langle x^2, \boldsymbol{q}_2 \rangle \, \boldsymbol{q}_2 = \left(x - \frac{1}{2}\right)$. Also, $\langle x^2, \boldsymbol{q}_1 \rangle = \int_0^1 x^2 \, dx = \frac{1}{3}$, so $\boldsymbol{u}_3 = x^2 - \left(x - \frac{1}{2}\right) - \frac{1}{3} = x^2 - x + \frac{1}{6}$.

We have $||\boldsymbol{u}_3||^2 = \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 \, dx = \int_0^1 x^4 - 2x^3 + \frac{4}{3}x^2 - \frac{1}{3}x + \frac{1}{36} \, dx = \frac{1}{180}$ and so $\boldsymbol{q}_3 = \sqrt{5}(6x^2 - 6x + 1)$.

A SageCell page that does computations

5. Let

(a)

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix}$$

Let

$$\begin{aligned}
\boldsymbol{q}_1 &= \boldsymbol{a}_1 = (2,1,2)^T \\
\boldsymbol{q}_2 &= \boldsymbol{a}_2 - \frac{\langle \boldsymbol{a}_1, \boldsymbol{q}_1 \rangle}{\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle} \boldsymbol{q}_1 \\
&= (1,1,1)^T - \frac{\langle (2,1,2)^T, (1,1,1)^T \rangle}{\langle (2,1,2)^T, (2,1,2)^T \rangle} (2,1,2)^T \\
&= (1/9)(-1,4,-1)^T
\end{aligned}$$

Check that this is orthogonal to $(2,1,2)^T$.

Now just normalize these

$$\hat{\boldsymbol{q}}_{\mathbf{1}} = (2/3, 1/3, 2/3)^T$$
$$\hat{\boldsymbol{q}}_{\mathbf{2}} = (1/(3\sqrt{2}))(-1, 4, -1)^T$$

So

$$Q = \begin{bmatrix} 2/3 & -1/(3\sqrt{2}) \\ 1/3 & 4/(3\sqrt{2}) \\ 2/3 & -1/(3\sqrt{2}) \end{bmatrix}$$

$$R = Q^T A = \begin{bmatrix} \langle \hat{\boldsymbol{q}}_{\mathbf{1}}, \boldsymbol{a}_1 \rangle & \langle \hat{\boldsymbol{q}}_{\mathbf{1}}, \boldsymbol{a}_2 \rangle \\ 0 & \langle \hat{\boldsymbol{q}}_{\mathbf{2}}, \boldsymbol{a}_2 \rangle \end{bmatrix} = \begin{bmatrix} 3 & 5/3 \\ 0 & 2/(3\sqrt{2}) \end{bmatrix}.$$

Check: $A = QR$.

7. We know $\mathrm{CS}(A) \perp \mathrm{NS}(A)$ where $A = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 3 & 5 \end{bmatrix}$. To find $\ker(A)$ start with $\mathrm{rref}(A) = \begin{bmatrix} 1 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$. So letting $x_2 = s$ and $x_4 = t$ we have

$$\boldsymbol{x} \in \ker(A) \Longleftrightarrow \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 4 \\ 0 \\ -3 \\ 1 \end{bmatrix}$$

so a basis for $\ker(A) = \mathrm{span}\{(-1, 1, 0, 0)^T, (4, 0, -3, 1)^T\}$

Check that these are indeed orthogonal to the given vectors.

Now use GS

$$\boldsymbol{q}_3 = (-1, 1, 0, 0)^T$$
$$\boldsymbol{q}_4 = (4, 0, -3, 1)^T - \frac{\langle (4, 0, -3, 1)^T, (-1, 1, 0, 0)^T \rangle}{\langle (-1, 1, 0, 0)^T, (-1, 1, 0, 0)^T \rangle}(-1, 1, 0, 0)^T = (2, 2, -3, 1)^T$$

Now just normalize to make these into unit vectors.

$$\hat{\boldsymbol{q}}_{\mathbf{3}} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \hat{\boldsymbol{q}}_{\mathbf{4}} = \frac{1}{3\sqrt{2}} \begin{bmatrix} 2 \\ 2 \\ -3 \\ 1 \end{bmatrix}$$

8. Use Gram-Schmidt to find orthonormal basis for $\mathrm{span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ where

$$\boldsymbol{x}_1 = \begin{bmatrix} 4 \\ 2 \\ 2 \\ 1 \end{bmatrix} \quad \boldsymbol{x}_2 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix} \quad \boldsymbol{x}_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

$$
\begin{aligned}
\boldsymbol{q}_1 &= \boldsymbol{x}_1 \\
\boldsymbol{q}_2 &= \boldsymbol{x}_2 - \frac{\langle \boldsymbol{q}_1, \boldsymbol{x}_2 \rangle}{\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle}\boldsymbol{q}_1 \\
&= (2,0,0,2) - \frac{10}{25}(4,2,2,1) \\
&= (2/5, -4/5, -4/5, 8/5) = (2/5)(1,-2,-2,4) \\
\boldsymbol{q}_3 &= \boldsymbol{x}_3 - \frac{\langle \boldsymbol{x}_3, \boldsymbol{q}_1 \rangle}{\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle}\boldsymbol{q}_1 - \frac{\langle \boldsymbol{x}_3, \boldsymbol{q}_2 \rangle}{\langle \boldsymbol{q}_2, \boldsymbol{q}_2 \rangle}\boldsymbol{q}_2 \\
&= (1,1,-1,1) - \frac{5}{25}(4,2,2,1) - \frac{(2/5)(5)}{(2/5)^2(25)}(2/5)(1,-2,-2,4) \\
&= (1,1,-1,1) - (4/5, 2/5, 2/5, 1/5) - (1/5, -2/5, -2/5, 4/5) \\
&= (0,1,-1,0) = (0,1,-1,0)
\end{aligned}
$$

So the final normalized orthonormal basis is

$$
\boldsymbol{q}_1 = \left(\frac{1}{5}\right)\begin{bmatrix} 4 \\ 2 \\ 2 \\ 1 \end{bmatrix} \quad \boldsymbol{q}_2 = \left(\frac{1}{5}\right)\begin{bmatrix} 1 \\ -2 \\ -2 \\ 4 \end{bmatrix} \quad \boldsymbol{q}_3 = \left(\frac{1}{2}\right)^{1/2}\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}
$$

9. The modified Gram-Schmidt looks like:

First pass:

$$
\begin{aligned}
\boldsymbol{q}_1 &= \boldsymbol{x}_1 \\
\boldsymbol{q}_2 &= \boldsymbol{x}_2 - \frac{\langle \boldsymbol{q}_1, \boldsymbol{x}_2 \rangle}{\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle}\boldsymbol{q}_1 \\
&= (2,0,0,2) - \frac{10}{25}(4,2,2,1) \\
&= (2/5, -4/5, -4/5, 8/5) = (2/5)(1,-2,-2,4) \\
\boldsymbol{q}_3 &= \boldsymbol{x}_3 - \frac{\langle \boldsymbol{x}_3, \boldsymbol{q}_1 \rangle}{\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle}\boldsymbol{q}_1 \\
&= (1,1,-1,1) - \frac{5}{25}(4,2,2,1) \\
&= (1/5, 3/5, -7/5, 4/5) = (1/5)(1,3,-7,4)
\end{aligned}
$$

Now $\boldsymbol{q}_2$ and $\boldsymbol{q}_3$ are orthogonal to $\boldsymbol{q}_1$.

Second pass:

$$
\begin{aligned}
\boldsymbol{q}_1' &= \boldsymbol{q}_1 \\
\boldsymbol{q}_2' &= \boldsymbol{q}_2 = (2/5)(1,-2,-2,4) \\
\boldsymbol{q}_3' &= \boldsymbol{q}_3 - \frac{\langle \boldsymbol{q}_3, \boldsymbol{q}_2 \rangle}{\langle \boldsymbol{q}_2, \boldsymbol{q}_2 \rangle}\boldsymbol{q}_2 \\
&= (1/5)(1,3,-7,4) - \frac{(1/5)(2/5)25}{(2/5)^2(25)}(2/5)(1,-2,-2,4) \\
&= (1/5)(1,3,-7,4) - (1/5)(1,-2,-2,4) \\
&= (1/5)(0,5,-5,0) = (0,1,-1,0)
\end{aligned}
$$

Now just normalize and get the same answer as in (8).

14. and 15. Let $\mathcal{B}$ be a basis for $W = U \cap V$ and let $\mathcal{B}_U \supseteq \mathcal{B}$ be a basis for U that extends $\mathcal{B}_W$. Extend $\mathcal{B}_U$ to $\mathcal{B}_{V+U}$ a basis for $V + U$. For each $\boldsymbol{z} \in \mathcal{B}_{V+U} - \lfloor_U$, $\boldsymbol{z} = \boldsymbol{v}_z + \boldsymbol{u}_z$ and as $\boldsymbol{z} \notin U$ we know $\boldsymbol{v}_z \neq 0$ and if $\mathcal{C} = \langle \boldsymbol{w}_z \mid \boldsymbol{z} \in \mathcal{B}_{V+U} - \lfloor_U \rangle$, then clearly $\mathcal{B}_U \cup \mathcal{C}$ is still independent and spans $V + U$. So we may assume $\mathcal{B}_{U+V} - \mathcal{B}_U \subset V$.

Claim: $\mathcal{B}_V \stackrel{\text{def}}{=} \mathcal{B}_W \cup (\mathcal{B}_{U+V} - \mathcal{B}_U)$ is a basis for V.

Independence is for free since $\mathcal{B}_{V+U}$ is independent. So we must show that $\mathcal{B}_V$ spans V. Let $\boldsymbol{v} \in V$ and $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{u} + \boldsymbol{z}$ where $\boldsymbol{w} \in \operatorname{span}(\mathcal{B}_W)$, $\boldsymbol{u} \in \operatorname{span}(\mathcal{B}_U - \mathcal{B}_W)$, and $\boldsymbol{z} \in \operatorname{span}(\mathcal{B}_{U+V} - \mathcal{B}_U) \subseteq V$.

If $\boldsymbol{u} \neq \boldsymbol{0}$, then $\boldsymbol{u} = \boldsymbol{v} - (\boldsymbol{w} + \boldsymbol{z}) \in V$, but then $\boldsymbol{u} \in U \cap V = W$ which is impossible since then a non-zero linear combination from $\mathcal{B}_U - \mathcal{B}_W$ is also a linear combination from $\mathcal{B}_W$, contradicting the independence of $\mathcal{B}_U$.

So $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{z} \in \operatorname{span}(\mathcal{B}_V)$. The

This shows $|\mathcal{B}_U \cup \mathcal{B}_V| = |\mathcal{B}_U| + |\mathcal{B}_V| - |\mathcal{B}_U \cap \mathcal{B}_V|$ so $\dim(U+V) = \dim(U) + \dim(V) - \dim(U \cap V)$.