MTH5112 Linear Algebra I
MTH5212 Applied Linear Algebra

COURSEWORK 10 — SOLUTIONS

Exercise (∗) 1. The solutions will appear on WeBWork after CW10 due date.

Exercise 2. (a) We know that H has an orthogonal basis because we can always construct an orthogonal basis by starting with an arbitrary basis and using the Gram–Schmidt process. In particular, we can, if we like, define $\hat{\mathbf{y}}$ by setting

(1) $$\hat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\right)\mathbf{v}_1 + \cdots + \left(\frac{\mathbf{y}\cdot\mathbf{v}_r}{\mathbf{v}_r\cdot\mathbf{v}_r}\right)\mathbf{v}_r.$$

(b) We now define $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Then $\mathbf{z}$ is orthogonal to all of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ in the chosen orthogonal basis for H, because for each $i \in \{1, \ldots, r\}$ we have

$$\begin{aligned}
\mathbf{z}\cdot\mathbf{v}_i &= (\mathbf{y} - \hat{\mathbf{y}})\cdot\mathbf{v}_i \\
&= \left(\mathbf{y} - \left(\frac{\mathbf{y}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\right)\mathbf{v}_1 - \cdots - \left(\frac{\mathbf{y}\cdot\mathbf{v}_r}{\mathbf{v}_r\cdot\mathbf{v}_r}\right)\mathbf{v}_r\right)\cdot\mathbf{v}_i \\
&= \mathbf{y}\cdot\mathbf{v}_i - \left(\frac{\mathbf{y}\cdot\mathbf{v}_i}{\mathbf{v}_i\cdot\mathbf{v}_i}\right)\mathbf{v}_i\cdot\mathbf{v}_i \qquad \text{(because } \mathbf{v}_i\cdot\mathbf{v}_j = 0 \text{ for all } j \neq i\text{)} \\
&= \mathbf{y}\cdot\mathbf{v}_i - \mathbf{y}\cdot\mathbf{v}_i \\
&= 0.
\end{aligned}$$

Hence, it follows from Theorem 6.8(ii) in lectures that $\mathbf{z}$ is orthogonal to *every* vector in H, i.e. $\mathbf{z} \in H^\perp$.

(c) Suppose, as suggested, that

$$\mathbf{y} = \hat{\mathbf{y}}_1 + \mathbf{z}_1 \quad \text{for some} \quad \hat{\mathbf{y}}_1 \in H \text{ and } \mathbf{z}_1 \in H^\perp.$$

Then we have

$$\hat{\mathbf{y}}_1 + \mathbf{z}_1 = \hat{\mathbf{y}} + \mathbf{z},$$

because both sides of this equation are equal to $\mathbf{y}$, and so

(2) $$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z} - \mathbf{z}_1.$$

This implies that the vector $\hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ is an element of *both* H and $H^\perp$, i.e. it is in H because both $\hat{\mathbf{y}}$ and $\hat{\mathbf{y}}_1$ are in H, and it is in $H^\perp$ because $\mathbf{z} - \mathbf{z}_1$ is in $H^\perp$ (because $H^\perp$ is a subspace). In particular, $\hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ must be orthogonal to itself, and so

$$(\hat{\mathbf{y}} - \hat{\mathbf{y}}_1)\cdot(\hat{\mathbf{y}} - \hat{\mathbf{y}}_1) = 0, \quad \text{i.e.} \quad ||\hat{\mathbf{y}} - \hat{\mathbf{y}}_1|| = 0.$$

Therefore, $\hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ is the zero vector, so $\hat{\mathbf{y}}_1 = \hat{\mathbf{y}}$, and hence by (2) we also have $\hat{\mathbf{z}}_1 = \hat{\mathbf{z}}$.

Exercise 3. If $\mathbf{x} \in H$ then $\mathbf{x}$ is orthogonal to every vector in $H^\perp$ (because every vector in $H^\perp$ is orthogonal to $\mathbf{x}$) and so $\mathbf{x} \in (H^\perp)^\perp$. Therefore, H is a subset of $(H^\perp)^\perp$. It remains to show that $(H^\perp)^\perp$ is a subset of H. Suppose that $\mathbf{x} \in (H^\perp)^\perp$. By the Orthogonal Decomposition Theorem, we can write $\mathbf{x}$ in the form $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{z}$ where $\hat{\mathbf{x}} \in H$ and $\mathbf{z} \in H^\perp$. Since $\mathbf{z} \in H^\perp$, it is orthogonal to both $\mathbf{x}$ (which is an element of H) and $\hat{\mathbf{x}}$ (which is an element of $(H^\perp)^\perp$). Therefore,

$$\mathbf{z} \cdot \mathbf{z} = \mathbf{z} \cdot \mathbf{z} + \underbrace{\mathbf{z} \cdot \hat{\mathbf{x}}}_{=0} = \mathbf{z} \cdot (\mathbf{z} + \hat{\mathbf{x}}) = \mathbf{z} \cdot \mathbf{x} = 0.$$

Hence, $\mathbf{z} = 0$, and so $\mathbf{x} = \hat{\mathbf{x}} \in H$. Therefore, $(H^\perp)^\perp$ is a subset of H.

Exercise 4. (a) We have

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\right) \mathbf{u} = \frac{-1 \cdot 1 + 7 \cdot 3}{1 \cdot 1 + 3 \cdot 3}(1,3)^T = 2(1,3)^T = (2,6)^T.$$

(b) We have

$$\mathbf{y} - \hat{\mathbf{y}} = (-1,7)^T - (2,6)^T = (-3,1)^T \quad \text{and hence} \quad (\mathbf{y} - \hat{\mathbf{y}}) \cdot \mathbf{u} = (-3,1)^T \cdot (1,3)^T = 0.$$

(c) Geometrically, $\mathbf{y} - \hat{\mathbf{y}}$ is perpendicular to $\mathbf{u}$, and $||\mathbf{y} - \hat{\mathbf{y}}||$ is the distance from $\hat{\mathbf{y}}$ to $\mathbf{y}$. Hence, $||\mathbf{y} - \hat{\mathbf{y}}||$ is the shortest distance from $\mathbf{y}$ to the line through the origin in the direction of $\mathbf{u}$.

Exercise 5. We have

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = (1,2,1)^T \cdot (-3,1,1)^T = 1 \cdot (-3) + 2 \cdot 1 + 1 \cdot 1 = -3 + 2 + 1 = 0,$$

so $\{\mathbf{u}_1, \mathbf{u}_2\}$ is certainly an orthogonal set, hence a linearly independent set, and therefore a basis for $H = \text{span}(\mathbf{u}_1, \mathbf{u}_2)$. Using the Orthogonal Decomposition Theorem (see Exercise 1), we find that the orthogonal decomposition of the given vector $\mathbf{y} = (6,-1,8)^T$ is

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z},$$

where

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1 + \left(\frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right) \mathbf{u}_2 = \frac{12}{6}(1,2,1)^T + \frac{-11}{11}(-3,1,1)^T = (5,3,1)^T$$

and

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = (6,-1,8)^T - (5,3,1)^T = (1,-4,7)^T.$$

Exercise 6.

(a) To see that the set $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is linearly independent, observe that the matrix

$$\begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -2 \\ 0 & 0 & 1 \\ 1 & 1 & 4 \\ 0 & 1 & -3 \end{pmatrix} \quad \text{is row equivalent to} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

and hence has rank 3. The set $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is not orthogonal because, e.g. $\mathbf{x}_1 \cdot \mathbf{x}_2 = 4 \neq 0$.

(b) We use the Gram–Schmidt process to build an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for H from the given basis $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. We begin the by setting

$$\mathbf{v}_1 = \mathbf{x}_1 = (1,0,1,0)^T.$$

The next step is to define the vector $\mathbf{v}_2$ by subtracting the orthogonal projection of $\mathbf{x}_2$ onto span$(\mathbf{v}_1)$ from $\mathbf{x}_2$:

$$\mathbf{v}_2 = \mathbf{x}_2 - \left(\frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right) \mathbf{v}_1 = (3,0,1,1)^T - \frac{4}{2}(1,0,1,0)^T = (1,0,-1,1)^T.$$

Finally, $\mathbf{v}_3$ is defined by subtracting the orthogonal projection of $\mathbf{x}_3$ onto span$(\mathbf{v}_1, \mathbf{v}_2)$ from $\mathbf{x}_3$:

$$\begin{aligned}\mathbf{v}_3 &= \mathbf{x}_3 - \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 \\ &= (-2, 1, 4, -3)^T - \frac{2}{2}(1, 0, 1, 0)^T - \frac{-9}{3}(1, 0, -1, 1)^T \\ &= (0, 1, 0, 0)^T.\end{aligned}$$

(c) We have $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}} \in H$ and $\mathbf{z} \in H^\perp$ are given by

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 + \left(\frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 + \left(\frac{\mathbf{y} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3}\right)\mathbf{v}_3 = \frac{6}{2}\mathbf{v}_1 + \frac{-6}{2}\mathbf{v}_2 + \frac{1}{1}\mathbf{v}_3 = (1, 1, 5, -2)^T$$

and

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = (1, 0, -1, -2)^T.$$

By the Best Approximation Theorem (Corollary 6.12 in lectures), the best approximation to $\mathbf{y}$ by a vector in H is the vector $\hat{\mathbf{y}} = (1, 1, 5, -2)^T$.

Exercise 7. The normal equations are (by definition)

$$A^T A\mathbf{x} = A^T\mathbf{b},$$

so we first compute

$$A^T A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

and

$$A^T\mathbf{b} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 8 \end{pmatrix}.$$

The normal equations are therefore

$$\begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 2 \\ 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 8 \end{pmatrix}.$$

Using Gaussian elimination, you should find that the augmented matrix corresponding to this system is row equivalent to

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

From this we see that x_3 is a free variable, while x_1 and x_2 are leading variables. Writing $\alpha = x_3$ to emphasise that x_3 is a free variable, we find that $x_2 = 3 - \alpha$ and $x_1 = 1 - \alpha$. The set of least squares solutions of the original system $A\mathbf{x} = \mathbf{b}$ is therefore

$$\{(1 - \alpha, 3 - \alpha, \alpha)^T \mid \alpha \in \mathbb{R}\}.$$