1. Written Problems

Chapter 6.3

1. The vector in $\text{Span}\{u_4\}$ is

\[
\frac{x \cdot u_4}{u_4 \cdot u_4} u_4 = \frac{72}{36} u_4 = 2u_4 = \begin{bmatrix} -6 \\ -2 \\ 2 \end{bmatrix}
\]

Since $x = c_1 u_1 + c_2 u_2 + c_3 u_3 + \frac{x \cdot u_4}{u_4 \cdot u_4} u_4$, the vector

\[
x - \frac{x \cdot u_4}{u_4 \cdot u_4} u_4 = \begin{bmatrix} 10 \\ -8 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} -6 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 0 \\ 0 \end{bmatrix}
\]

is in $\text{Span}\{u_1, u_2, u_3\}$.

4. Since $u_1 \cdot u_2 = -12 + 12 + 0 = 0$, $\{u_1, u_2\}$ is an orthogonal set. The orthogonal projection of $y$ onto $\text{Span}\{u_1, u_2\}$ is

\[
\hat{y} = y \cdot u_1 \frac{u_1}{u_1 \cdot u_1} + y \cdot u_2 \frac{u_2}{u_2 \cdot u_2} = \frac{30}{25} u_1 - \frac{15}{25} u_2 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} - \begin{bmatrix} -4 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}
\]

11. Note that $v_1$ and $v_2$ are orthogonal. The Best Approximation Theorem says that $\hat{y}$, which is the orthogonal projection of $y$ onto $W = \text{Span}\{v_1, v_2\}$, is the closest point to $y$ in $W$. This vector is

\[
\hat{y} = \frac{y \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{y \cdot v_2}{v_2 \cdot v_2} v_2 = \frac{1}{2} v_1 + \frac{3}{2} v_2 = \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}
\]
15. The distance from the point \( y \) in \( \mathbb{R}^3 \) to a subspace \( W \) is defined as the distance from \( y \) to the closest point in \( W \). Since the closest point in \( W \) to \( y \) is \( \hat{y} = \text{proj}_W y \), the desired distance is \( \| y - \hat{y} \| \). One computes that \( \hat{y} = \begin{bmatrix} 3 \\ -9 \\ -1 \end{bmatrix} \), \( y = \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix} \), and \( \| y - \hat{y} \| = \sqrt{40} = 2\sqrt{10} \).

22. a. True. See the proof of the Orthogonal Decomposition Theorem.
   b. True. See the subsection "A Geometric Interpretation of the Orthogonal Projection."
   c. True. The orthogonal decomposition in Theorem 8 is unique.
   d. False. The Best Approximation Theorem says that the best approximation to \( y \) is \( \text{proj}_W y \).
   e. False. This statement is only true if \( x \) is in the column space of \( U \). If \( n > p \), then the column space of \( U \) will not be all of \( \mathbb{R}^n \), so the statement cannot be true for all \( x \) in \( \mathbb{R}^n \).

Chapter 6.4

2. Set \( v_1 = x_1 \) and compute that \( v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = x_2 - \frac{1}{2} v_1 = \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix} \). Thus an orthogonal basis for \( W \) is \( \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix} \).

5. Set \( v_1 = x_1 \) and compute that \( v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = x_2 - 2v_1 = \begin{bmatrix} 5 \\ 1 \\ -4 \end{bmatrix} \). Thus an orthogonal basis for \( W \) is \( \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -4 \end{bmatrix} \).
10. Call the columns of the matrix $x_1$, $x_2$, and $x_3$ and perform the Gram-Schmidt process on these vectors:

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = x_2 - (-3) v_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 = x_3 - \frac{1}{2} v_1 - \frac{5}{2} v_2 = \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$$

Thus an orthogonal basis for $W$ is

$$\begin{bmatrix} -1 & 3 & -1 \\ 3 & 1 & -1 \\ 1 & 1 & 3 \end{bmatrix}.$$

13. Since $A$ and $Q$ are given,

$$R = Q^T A = \begin{bmatrix} 5/6 & 1/6 & -3/6 & 1/6 \\ -1/6 & 5/6 & 1/6 & 3/6 \\ -3 & -5 & 6 & 12 \\ 1 & 5 \end{bmatrix}.$$

17. a. False. Scaling was used in Example 2, but the scale factor was nonzero.
b. True. See (1) in the statement of Theorem 11.
c. True. See the solution of Example 4.

Chapter 6.5
3. To find the normal equations and to find $\hat{x}$, compute

$$A' A = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & 5 \\ 0 & 2 & 5 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix}$$

$$A' b = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & 5 \\ 0 & 2 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -4 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

a. The normal equations are $(A^T A)x = A^T b$:

$$\begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

b. Compute

$$\hat{x} = (A^T A)^{-1} A^T b = \frac{1}{216} \begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ -6 \end{bmatrix} = \begin{bmatrix} 4/3 \\ -1/3 \end{bmatrix}$$

5. To find the least squares solutions to $Ax = b$, compute and row reduce the augmented matrix for the system $A^T A x = A^T b$:

$$\begin{bmatrix} A^T A & A^T b \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 & 14 \\ 2 & 2 & 0 & 4 \\ 2 & 0 & 2 & 10 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so all vectors of the form $\hat{x} = \begin{bmatrix} 5 \\ -3 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ are the least-squares solutions of $Ax = b$. 

4
7. From Exercise 3, \( A = \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix} \), \( b = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \), and \( \hat{x} = \begin{bmatrix} 4/3 \\ -1/3 \end{bmatrix} \). Since

\[
A\hat{x} - b = \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 4/3 \\ -1/3 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -4 \\ 2 \end{bmatrix} = -1
\]

the least squares error is \( \| A\hat{x} - b \| = \sqrt{20} = 2\sqrt{5} \).

12. (a) Because the columns \( a_1, a_2 \) and \( a_3 \) of \( A \) are orthogonal, the method of Example 4 may be used to find \( \hat{b} \), the orthogonal projection of \( b \) onto \( \text{Col} A \):

\[
\hat{b} = \frac{b \cdot a_1}{a_1 \cdot a_1} a_1 + \frac{b \cdot a_2}{a_2 \cdot a_2} a_2 + \frac{b \cdot a_3}{a_3 \cdot a_3} a_3 = \frac{1}{3} a_1 + \frac{14}{3} a_2 + \left( -\frac{5}{3} \right) a_3
\]

\[
= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 17 \\ 0 \\ 3 \end{bmatrix} - \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 12 \\ 2 \\ 3 \end{bmatrix}
\]

(b) The vector \( \hat{x} \) contains the weights which must be placed on \( a_1, a_2, \) and \( a_3 \) to produce \( \hat{b} \). These weights are easily read from the above equation, so \( \hat{x} = \begin{bmatrix} 1/3 \\ 14/3 \\ -5/3 \end{bmatrix} \).

18. a. True. See the paragraph following the definition of a least-squares solution.
   b. False. If \( \hat{x} \) is the least-squares solution, then \( A\hat{x} \) is the point in the column space of \( A \) closest to \( b \). See Figure 1 and the paragraph preceding it.
   c. True. See the discussion following equation (1).
   d. False. The formula applies only when the columns of \( A \) are linearly independent. See Theorem 14.
   e. False. See the comments after Example 4.
   f. False. See the Numerical Note.
2. MATLAB Problems

**NOTE:** The results/outputs for the matrices and the vectors are not going to be the same for everyone since we are dealing with random vectors.

**Problem 2.1**

By opening a script in MATLAB and typing the following:

```matlab
clear all; close all;
u = rand(2,1);
v = rand(2,1);
rk = rank([u, v])  % part a)

norm1 = norm(u) + norm(v)
norm2 = norm(u + v)  % part b)
w = (u'*v)/(v'*v)*v
z = u - w
orth = v'*z  % part c)
P = v*inv(v'*v)*v'
w1 = P*u
ww1 = w - w1
```

we can answer all the parts of problem 1.

a) By defining vectors $u$ and $v$ and checking if they are linearly independent (i.e. if the rank is 2), we can check the triangle inequality. We get $||u|| + ||v|| = 1.4086 \geq ||u + v|| = 1.2556$.

b) We have that $w = [0.4189 \ 0.1040]^T$, $z = w - u = [-0.1429 \ 0.5757]^T$
and we can see that $z$ is orthogonal to $v$ since $z \cdot v = 1.3878e-017 = 0$.

c) For matrix $P$ we have that $P =$

\[
\begin{bmatrix}
0.9420 & 0.2338 \\
0.2338 & 0.0580
\end{bmatrix}
\]

$P$ is a $2 \times 2$ matrix since $v$ is $2 \times 1$, $inv(v^Tv)$ is just the inverse of the $\text{norm}^2(v)$ which is a scalar quantity and $v' = v^T$ is a $1 \times 2$ vector. Then, $Pu = w$ since their difference $Pu - w = 1.0e-016 \ast [-0.5551 \quad -0.13880]^T = 0$

### Problem 2.2

By opening a new script in MATLAB and typing the following, we can answers Problem 2.2:

```matlab
clear all;
a1 = rand(5,1);
a2 = rand(5,1);
a3 = rand(5,1);
A = [a1, a2, a3];
rank(A)
%a)
[Q,R] = qr(A,0)

QtQ = Q' * Q
v1 = (1/norm(a1))*a1; % normalize first column vector
v2 = a2 - ((a2'*v1)/(v1'*v1))*v1; % perpendicular component of projection
v2 = (1/norm(v2))*v2; % normalize second vector
v3 = a3 - ((a3'*v1)/(v1'*v1))*v1 - ((a3'*v2)/(v2'*v2))*v2;

Q = [v1, v2, v3] % should be 5x3 matrix
Qcheck = Q' * Q % should be 3x3 identity matrix
R = Q' * A % should be 3x3 upper triangular

%b)
```

7
P = A * inv(A' * A) * A'

b = rand(5, 1);
w = P * b

% c)
P1 = Q * Q'
diff = P - P1

% d)
w' * (b - w)

We check again that $a_1$, $a_2$ and $a_3$ are linearly dependent by looking at the rank of matrix $A$ (which has as columns the vectors $a_1, a_2, a_3$). It must be that $\text{rank}(A) = 3$. 

a) We can do the $QR$ factorization either by using the $qr(A, 0)$ command or by creating the columns of $Q$ and done in the book, either way we will have the same answer. We get that $Q =$

$$
\begin{bmatrix}
-0.0929 & -0.1894 & 0.7464 \\
-0.3889 & -0.5320 & 0.3674 \\
-0.7490 & 0.6491 & 0.1014 \\
-0.2656 & -0.3533 & -0.2803 \\
-0.4567 & -0.3675 & -0.4679
\end{bmatrix}
$$

and $R =$

$$
\begin{bmatrix}
-1.2814 & -0.9577 & -0.9707 \\
0 & -0.7121 & -0.4277 \\
0 & 0 & 0.9642
\end{bmatrix}
$$

Also, $Q^T Q$ should be a $3 \times 3$ identity matrix, no matter what your $Q$ and $R$ matrices are. This shows that the columns of $Q$ are orthonormal, i.e. orthogonal to each other and of unit length. $Q^T Q =$

$$
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
$$

which means that $v_1^T v_1 = 1$, $v_1^T v_2 = 0$, $v_1^T v_3 = 0$ etc, showing that the columns of $Q$ are of unit length and orthogonal to each other since the dot product is zero.
b) Matrix $P$ is found to be

$$
P = \begin{bmatrix}
0.6017 & 0.4111 & 0.0223 & -0.1177 & -0.2373 \\
0.4111 & 0.5692 & -0.0167 & 0.1882 & 0.2012 \\
0.0223 & -0.0167 & 0.9926 & -0.0588 & 0.0561 \\
-0.1177 & 0.1882 & -0.0588 & 0.2739 & 0.3823 \\
-0.2373 & 0.2012 & 0.0561 & 0.3823 & 0.5626 
\end{bmatrix}
$$

and the projection $w$ of a random vector $b$ onto the column space of $A$ is

$$
w = Pb = [0.3527 \ 0.7824 \ 0.2099 \ 0.4292 \ 0.5706]^T.
$$

We can more easily calculate $P$ as $P = QQ^T$. We can see that we get the same $P$s in part b) by calculating their difference and seeing that $QQ^T - P$ is a $5 \times 5$ zero matrix.

\textbf{Proof}:

\begin{align*}
A &= QR \\
P &= A(A^TA)^{-1}A^T = QR((QR)^T(QR))^{-1}(QR)^T \\
P &= QR(R^TQ^TQR)^{-1}R^TQ^T \\
Q^TQ &= I \\
P &= QR(R^TR)^{-1}R^TQ^T \\
P &= QRR^{-1}(R^T)^{-1}R^TQ^T = QIQ^T \Rightarrow \\
\Rightarrow P &= QQ^T
\end{align*}

d) Finally $b$ can be broken into two orthogonal vector $w$ and $b - w$ since $w \cdot (b - w) = 0$

\textbf{Problem 2.3}

By typing the following in MATLAB (in the previous script below problem 2.2) we have :

a)  
\texttt{nosol=rref([A,b])}
%b)  
\texttt{sol_b=rref([A,w])}
%c)  
\texttt{y=rand(3,1);}  
\texttt{PAy=P*A*y}  
\texttt{Ay=A*y}
a) The equation $Ax = b$ does not have a solution since by rrefing the augmented matrix $rref([A, b])$ we have that

$$
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}
$$

We can see the inconsistency in row 4 since it shown that $0 = 1$, so that means that there is no solution to the system.

We can see that by doing $rref([A, w])$ we indeed get a solution for the system $A\tilde{x} = w$ i.e.

$$
\begin{bmatrix}
1.0000 & 0 & 0 & -0.1243 \\
0 & 1.0000 & 0 & 0.8794 \\
0 & 0 & 1.0000 & 0.1915 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
$$

c) We can verify that $PAy = Ay$ by calculating their difference and checking if it is zero, i.e. $PAy - Ay = 1.0e-015 \begin{bmatrix}
-0.4441 & 0 & -0.4441 & 0.1110 & 0.6661
\end{bmatrix}^T$ which is numerically the zero vector.

d) Finally, we can see that the projection $w$ of vector $b$ onto the column space of $A$, is closer to $b$ than any other random vector $y$ by showing that $||Ay - b|| > ||A\tilde{x} - b|| = ||w - b||$. Indeed, $||Ay - b|| = 1.0626 > ||w - b|| = 0.5193$.

GOOD LUCK ON THE FINAL AND HAPPY HOLIDAYS!