18.06 Problem Set 5

SOLUTIONS TO SELECTED PROBLEMS

1. Section 4.2, Problem 5

Answers: \( P_1 = \frac{1}{9} \begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix} \); \( P_2 = \frac{1}{9} \begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix} \); \( P_1 P_2 \) is the zero matrix because \( a_1 \perp a_2 \) (projecting a vector onto \( a_2 \) and then projecting the result onto \( a_1 \) gives 0).

2. Section 4.2, Problem 13

Solution: The column space of \( A \) is the “\( xyz \)-hyperplane” in the 4-dimensional space, so the projection of \( b = (1, 2, 3, 4) \) is \((1, 2, 3, 0)\). The projection matrix is \( P = \begin{bmatrix} I_3 & 0 \\ 0 & 0 \end{bmatrix} \); it is a 4 by 4 matrix.

3. Section 4.2, Problem 17

Solution: We compute: \((I - P)^2 = I^2 - IP - PI + P^2 = I - P - P + P = I - P\). If \( P \) projects onto the column space of \( A \), then \((I - P)v = v - Pv\) is the difference between a vector \( v \) and its projection onto the column space of \( A \), which is the projection of \( v \) onto the space \( A^\perp \), or the left nullspace of \( A \).

4. Section 4.2, Problem 19

Answers: For example, we can choose \( A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \). Then \( P = A(A^T A)^{-1} A^T = \frac{1}{6} \begin{bmatrix} 5 & 1 & 2 \\ 1 & 5 & -2 \\ 2 & -2 & 2 \end{bmatrix} \). (Of course, \( P \) will be the same for any choice of \( A \).)

5. Section 4.2, Problem 27

Solution: If \( A^T Ax = 0 \), then \( Ax = 0 \). The vector \( Ax \) is in the nullspace of \( A^T \), and \( Ax \) is always in the column space of \( A \). Since the nullspace of \( A^T \) is orthogonal to the column space of \( A \), in order to be in both of these spaces \( Ax \) has to be 0.

6. Section 4.3, Problem 1
Answers:  
\[ A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix} \]

\( A^T A = \begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix} \) and \( A^T b = \begin{bmatrix} 36 \\ 112 \end{bmatrix} \). 

\( A^T A \hat{x} = A^T b \) gives  
\[ \hat{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \]

and  
\[ p = A \hat{x} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix} \]

and  
\( e = b - p = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 3 \end{bmatrix} \). Then  
\[ E = \|e\|^2 = 44. \]

7. Section 4.3, Problem 17

Answers:  
\[ \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \\ 21 \end{bmatrix} \]

The solution  
\[ \hat{x} = \begin{bmatrix} 9 \\ 4 \end{bmatrix} \]

comes from  
\[ \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 35 \\ 42 \end{bmatrix} \].

8. Section 4.3, Problem 22

Answer: The least squares equation is  
\[ \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 5 \\ -10 \end{bmatrix} \]. Solution:  
\[ C = 1, D = -1. \]

9. Section 4.3, Problem 26

Solution: Equating the slopes of the line connecting \((t_1, b_1)\) and \((t_2, b_2)\) and the line connecting \((t_2, b_2)\) and \((t_3, b_3)\), we get the equation \((b_2 - b_1)/(t_2 - t_1) = (b_3 - b_2)/(t_3 - t_2)\). Another way to state the condition of \((t_i, b_i)\) being on one line is saying that \((b_1, b_2, b_3)\) is in the column space of the matrix  
\[ A = \begin{bmatrix} 1 & t_1 \\ 2 & t_2 \\ 3 & t_3 \end{bmatrix} \]. Equivalently, \((b_1, b_2, b_3)\) is orthogonal to the complimentary space to the column space of \(A\), which is spanned by the single vector  
\[ y = (t_2 - t_3, t_3 - t_1, t_1 - t_2) \]. Writing this condition algebraically, we get  
\[ (b_1, b_2, b_3) \cdot (t_2 - t_3, t_3 - t_1, t_1 - t_2) = 0, \]

or  
\[ (b_2 - b_1)(t_3 - t_2) = (b_3 - b_2)(t_2 - t_1) \] — essentially the same equation we got by equating slopes.

10. Section 4.3, Problem 27

Solution: The unsolvable system is  
\[ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix} \]. Then  
\[ A^T A = \]
\[
\begin{bmatrix}
4 & 0 & 0 \\
0 & 2 & 0 \\
0 & 0 & 2
\end{bmatrix}
\quad \text{and} \quad
A^T b = \begin{bmatrix}
8 \\
-3 \\
-3
\end{bmatrix}.
\]
Solving \( A^T A \hat{x} = A^T b \) yields \( \hat{x} = \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 2 \\ -3/2 \\ -3/2 \end{bmatrix} \).

At \((x, y) = (0, 0)\), the plane \( z = 2 - \frac{3}{2}x - \frac{3}{2}y \) has height 2, which is the average of 0,1,3,4.

11. Section 4.4, Problem 6

**Solution:** Unfortunately, the statement in the problem is false. The claim is true for *orthonormal* matrices: indeed, if \( Q_1^T Q_1 = I \) and \( Q_2^T Q_2 = I \), then \( (Q_1 Q_2)^T Q_1 Q_2 = Q_2^T (Q_1^T Q_1) Q_2 = Q_2^T Q_2 = I \). Note that if \( Q_1 \) is orthogonal and \( Q_2 \) is orthonormal, then \( Q_1 Q_2 \) is orthogonal.

12. Section 4.4, Problem 7

**Solution:** The least squares solution is the solution to \( Q^T Q \hat{x} = Q^T b \), which in this case reduces to \( \hat{x} = Q^T b \).

13. (a) Minimize
\[
\int_0^1 (c + dt - t^2)^2 dt = \int_0^1 (c^2 + 2cd + d^2 t - 2cd^2 - 2dt^3 + t^4) dt
\]
\[
= c^2 + cd + \frac{1}{3}d^2 - \frac{2}{3}c - \frac{2}{4}d + \frac{1}{5}
\]

\[c\text{-derivative:} \quad 2c + d = \frac{2}{3}
\]
\[d\text{-derivative:} \quad c + \frac{2}{3}d = \frac{2}{4}
\]

Solution: \( c = -\frac{1}{6} \) and \( d = 1 \): Best line \( y = t - \frac{1}{6} \).

Note: Dividing by 2 shows the 2 by 2 Hilbert matrix with \( h_{ij} = 1/(i + j - 1) \):

\[
\text{hilb}(2) = \begin{bmatrix}
1 & 1/2 \\
1/2 & 1/3
\end{bmatrix}
\]

(b) The 10 by 2 matrix is \( A = \begin{bmatrix} \text{ones}(10, 1) & (1 : 10)'/10 \end{bmatrix} \) and the column vector is \( b = (1 : 10)' \ast (1 : 10)'/100 \).

\[
A^T A \begin{bmatrix} C \\ D \end{bmatrix} = A^T b \quad \text{is} \quad \begin{bmatrix} 10 & 5.5 \\ 5.5 & 3.85 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 3.85 \\ 3.02 \end{bmatrix}
\]

\[
giving \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} -0.22 \\ 0.11 \end{bmatrix}.
\]
(c) The same calculation with 10 changed to 20 (and 100 to 400) comes closer to $c = -\frac{1}{6}, d = 1$:

$$A^T A \begin{bmatrix} C \\ D \end{bmatrix} = A^T b \quad \text{is} \quad \begin{bmatrix} 20 & 10.5 \\ 10.5 & 7.175 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7.175 \\ 5.5125 \end{bmatrix}$$

giving \( \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} -1.925 \\ 1.0500 \end{bmatrix} \).

The error in comparing $D$ to $d = 1$ dropped from .1 to .05 (exactly in half). The error in comparing $C$ to $c = -\frac{1}{6}$ dropped from $c - C = .0533$ to $c - C = .0258$ (nearly in half).