1. (a) No one.

(b) I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

Yao Cai

3. (a) $M = uv^T = \begin{pmatrix} 1 \\ 2 \end{pmatrix} (2 \;\; 3) = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}$

$Mx = \lambda x \Rightarrow (M - \lambda I)x = 0$ has nontrivial solution only when

$\det|M - \lambda I| = 0$

$$\begin{vmatrix} 2-\lambda & 3 \\ 4 & 6-\lambda \end{vmatrix} = 0$$

$(2-\lambda)(6-\lambda) - 12 = 0$

$\lambda^2 - 8\lambda = 0 \quad \Rightarrow \lambda = 0$ or $\lambda = 8$

when $\lambda = 0$:

$$\begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

$\Rightarrow \quad 2x_1 + 3x_2 = 0 \quad \Rightarrow \quad x_2 = \frac{-2}{3} x_1$

$x = \begin{pmatrix} x_1 \\ -\frac{2}{3}x_1 \end{pmatrix}$ with $x_1 \neq 0$.

when $\lambda = 8$

$$\begin{pmatrix} -6 & 3 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

$\Rightarrow 4x_1 - 2x_2 = 0 \Rightarrow x_2 = 2x_1$

$\Rightarrow x = \begin{pmatrix} x_1 \\ 2x_1 \end{pmatrix}$ with $x_1 \neq 0$.

the eigenvalues for M is 0 and 8 with corresponding eigenvectors $\begin{pmatrix} x_1 \\ -\frac{2}{3}x_1 \end{pmatrix}$ and $\begin{pmatrix} x_1 \\ 2x_1 \end{pmatrix}$ where $x_1 \neq 0$

(b) $Mx = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1+3x_2 \\ 4x_1+6x_2 \end{pmatrix} = \begin{pmatrix} 2x_1+3x_2 \\ 2(2x_1+3x_2) \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot (2x_1+3x_2)$

so $\dim(Mx) = 1$

$\Rightarrow \text{rank}(M) = 1$

$\det(A) = 0 \cdot 8 = 0$

due to rank-nullity theorem,

the dimension of null space of M is:

$2 - \text{rank}(M) = 1$

(c) apparently $\text{rank}(pq^T) = 1$

then $\dim \text{nul}(pq^T) = d - \text{rank}(pq^T) = d-1 \geq 1$ (assumed ≥ 2)

- $\lambda = 0$ is one of the eigenvalues, since null space is not empty, and multiplicity of $\lambda = 0$ is $d-1$

  eigenvectors correspond to $\lambda = 0$ are all vectors in null space $\{x : pq^T x = 0\}$

  $\text{nul}(pq^T) = \text{col}(qp^T)^\perp$, the column space of qp^T is simply spanned by q, so geometrically $\text{null}(pq^T)$ is the hyperplane $\perp q$.

- $(pq^T)p = p(q^Tp) = (q^Tp)p$ (since q^Tp is scaler)

  so $\lambda = q^Tp$ is the other eigenvalue with eigenvector p

eigenvalue	eigenvectors
0	$q^\perp$
q^Tp	p

$\text{rank}(pq^T) = 1$ $\quad \det(pq^T) = 0 \cdot q^Tp = 0$

$\dim \text{null}(pq^T) = d-1$

4.(a) the true model is $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, so

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_i & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix} \quad \text{or } Aw = y$$

the corrupted y is $(y_1 \cdots y_i+\varepsilon_i \cdots y_n)^T$, then the model learned by OLS is

$$\begin{pmatrix} w_1^* \\ w_2^* \end{pmatrix} = (A^TA)^{-1}A^T \begin{pmatrix} y_1 \\ \vdots \\ y_i+\varepsilon_i \\ \vdots \\ y_n \end{pmatrix}$$

$$= (A^TA)^{-1}A^T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + (A^TA)^{-1}A^T \begin{pmatrix} 0 \\ \vdots \\ \varepsilon_i \\ \vdots \\ 0 \end{pmatrix}$$

$$= (A^TA)^{-1}A^T \cdot Aw + (A^TA)^{-1}A^T \begin{pmatrix} 0 \\ \vdots \\ \varepsilon_i \\ \vdots \\ 0 \end{pmatrix}$$

$$= w + (A^TA)^{-1}A^T \begin{pmatrix} 0 \\ \vdots \\ \varepsilon_i \\ \vdots \\ 0 \end{pmatrix}$$

$$(A^TA)\begin{pmatrix} w_1^*-w_1 \\ w_2^*-w_2 \end{pmatrix} = A^T \begin{pmatrix} 0 \\ \vdots \\ \varepsilon_i \\ \vdots \\ 0 \end{pmatrix} = \varepsilon_i \begin{pmatrix} x_i \\ 1 \end{pmatrix} \quad ①$$

$$A^TA = \begin{pmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} = \begin{pmatrix} \sum_j x_j^2 & \sum_j x_j \\ \sum_j x_j & n \end{pmatrix}$$

the left side of ① can be any 2D vector, but right side of ① can only be the vectors $\parallel \begin{pmatrix} x_i \\ 1 \end{pmatrix}$, so we $\boxed{\text{cannot}}$ always find a ε_i for any pair of (w_1^*, w_2^*)

the counter examples are (w_1^*, w_2^*) such that $A^TA\begin{pmatrix} w_1^*-w_1 \\ w_2^*-w_2 \end{pmatrix}$ not parallel to $\begin{pmatrix} x_i \\ 1 \end{pmatrix}$

(even if we can choose which i to corrupt, the right side doesn't span 2D space)

(b) if at most two ε_i can be used, then

$$A^T A \begin{pmatrix} w_1^* - w_1 \\ w_2^* - w_2 \end{pmatrix} = A^T \begin{pmatrix} 0 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_j \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} x_i \varepsilon_i + x_j \varepsilon_j \\ \varepsilon_i + \varepsilon_j \end{pmatrix} = \begin{pmatrix} x_i & x_j \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_i \\ \varepsilon_j \end{pmatrix} \quad ②$$

now the right hand side of ② can also be any 2D vector since $x_i \neq x_j$, so we $\boxed{\text{can}}$ always find the combination of ε_i and ε_j to make equation ② true

(c) ① a few outliners can easily change the model we learned by OLS

② data security is very important, the more data points an adversary can control, the more easily he can change the model we learn

5. (a) undergraduate level Linear Algebra and first half of graduate course Math 221 (Matrix computations); know what is SVD and several algorithms of matrix decomposition.

(b) never took an optimization course

(c) undergraduate level probability and a graduate level statistic course Stat 201B; know what is MLE and Bayesian statistics

(d) undergraduate vector calculus; know what is Stokes' theorem.

Programming: ① completed a small chess program and a simple git system with Java (in 61B)

② use python on daily basis