1. We can express $\boldsymbol{X}$ in terms of vectors:

$$\underset{3\times 4}{\boldsymbol{X}} = \begin{bmatrix} 1 & 7 & 0 & -3 \\ 0 & -2 & 1 & 2 \\ -1 & -5 & -1 & 1 \end{bmatrix} = \left[\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 7 \\ -2 \\ -5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} \right] = [\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \boldsymbol{x}_3,\ \boldsymbol{x}_4]$$

(a) Solution 1: To prove the linear dependence of columns of $\boldsymbol{X}$, find a non-zero vector $\boldsymbol{a} \in \mathbb{R}^4$, satisfying

$$\boldsymbol{X}\boldsymbol{a} = \sum_{i=1}^{4} a_i \boldsymbol{x}_i = \boldsymbol{0},$$

where a_i is ith element of $\boldsymbol{a}$. For example, $\boldsymbol{a}$ could be $(-4, 1, 0, 1)'$, since $\boldsymbol{X}\boldsymbol{a} = -4 * \boldsymbol{x}_1 + 1 * \boldsymbol{x}_2 + 0 * \boldsymbol{x}_3 + 1 * \boldsymbol{x}_4 = \boldsymbol{0}$.

A way to arrive at such solution is to search for a solution to the system of the equations:

$$\begin{cases} a_1 * 1 + a_2 * 7 + a_3 * 0 + a_4 * (-3) = 0 \\ a_1 * 0 + a_2 * (-2) * + a_3 * 1 + a_4 * 2 = 0 \\ a_1 * (-1) + a_2 * (-5) + a_3 * (-1) + a_4 * 1 = 0 \end{cases}$$

Then, you can obtain $a_1 = -4t - \frac{7}{2}s$, $a_2 = t + \frac{s}{2}$, $a_3 = s$, $a_4 = t$, where $s, t \in \mathbb{R}$.
(You will get $\boldsymbol{a}$ in Solution 1, by choosing $s = 0, t = 1$.)

(b) A row reduced echelon form of $\boldsymbol{X}$ is below.

$$\boldsymbol{X} = \begin{bmatrix} 1 & 7 & 0 & -3 \\ 0 & -2 & 1 & 2 \\ -1 & -5 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 7 & 0 & -3 \\ 0 & -2 & 1 & 2 \\ 0 & 2 & -1 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 7 & 0 & -3 \\ 0 & -2 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 7 & 0 & -3 \\ 0 & 1 & -0.5 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 3.5 & 4 \\ 0 & 1 & -0.5 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Because the row echelon form has two non-zero rows, we know that matrix $\boldsymbol{X}$ has two linearly independent row vectors. Thus, the rank of matrix $\boldsymbol{X}$ is 2.

(c) By slide 21 of set 1,
1. Note that the matrix $\boldsymbol{X}$ has rank 2 (by 1.(b)).
Find any $r \times r$ nonsingular submatrix of $\boldsymbol{X}$ where $r = rank(\boldsymbol{X}) = 2$. Call this matrix $\boldsymbol{W}$. For example, choose

$$\boldsymbol{W} = \begin{bmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

2. Compute $(\boldsymbol{W}^{-1})'$. Since $\boldsymbol{W}$ is the identity matrix,

$$(\boldsymbol{W}^{-1})' = \boldsymbol{W} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

3. Replace each element of $\boldsymbol{W}$ in $\boldsymbol{X}$ with the corresponding element of $(\boldsymbol{W}^{-1})'$. Then the corresponding matrix is

$$\boldsymbol{X} = \begin{bmatrix} 1 & 7 & 0 & -3 \\ 0 & -2 & 1 & 2 \\ -1 & -5 & -1 & 1 \end{bmatrix}.$$

4. Replace all other elements in $\boldsymbol{X}$ with zeros. The resulting matrix is

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

5. Transpose the resulting matrix to obtain $\boldsymbol{G}$, a generalized inverse for $\boldsymbol{X}$.

$$\boldsymbol{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(d) Use the `R` function `ginv` in the `MASS` package .

```
> X=matrix(c(1,7,0,-3,0,-2,1,2,-1,-5,-1,1),nrow=3, byrow=TRUE)
> MASS::ginv(X)
```

We can obtain a generalized inverse matrix

$$\boldsymbol{G}^* = \begin{bmatrix} -0.005089059 & 0.07888041 & -0.07379135 \\ 0.061068702 & 0.05343511 & -0.11450382 \\ -0.048346056 & 0.24936387 & -0.20101781 \\ -0.081424936 & 0.26208651 & -0.18066158 \end{bmatrix}.$$

Notice that the generalized inverses for a singular matrix (such as $\boldsymbol{X}$ in this problem) are not unique.

(e) $\boldsymbol{X}^*$ is 3×2 matrix with rank 2. ($rank(\boldsymbol{X}) = 2$ by (b)). Since $\boldsymbol{x}_1, \boldsymbol{x}_3$ are linearly independent and $\boldsymbol{x}_2, \boldsymbol{x}_4$ can be generated by the linear combinations of $\boldsymbol{x}_1, \boldsymbol{x}_3$ ($\boldsymbol{x}_2 = 7 * \boldsymbol{x}_1 - 2 * \boldsymbol{x}_3$, $\boldsymbol{x}_4 = -3 * \boldsymbol{x}_1 + 2 * \boldsymbol{x}_3$), $\mathcal{C}([\boldsymbol{c_1}, \boldsymbol{c_3}]) = \mathcal{C}([\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4])$. Thus,

$$\boldsymbol{X}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix}.$$

Note that you can pick any two linearly independent columns of $\boldsymbol{X}$ for the columns of $\boldsymbol{X}^*$.

(f) This column space is equivalent to the set of all vectors in $\mathbb{R}^3$ whose entries sum to zero.

Comment : I also accept the following as an answer. "All the three-dimensional vectors in the column space of $\boldsymbol{X}$ can be generated by the linear combination of $\boldsymbol{x}_1, \boldsymbol{x}_3$."

2. Consider an arbitrary list of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m$ where $\boldsymbol{a}_i \in \mathbb{R}^n$, for $i = 1, 2, ..., m$. WLOG, let $\boldsymbol{a}_1$=$\mathbf{0}$. If we take $c_1 = a$, $a \in \mathbb{R}\backslash\{0\}$ and $c_i = 0$, for $i = 2, 3, ..., m$, then

$$\sum_{i=1}^{m} c_i \boldsymbol{a}_i = \mathbf{0}.$$

Since $c_1, c_2, ..., c_m$ are not all zero, $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m$ are linearly dependent. Because $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m$ was an arbitrary list containing the zero vector, any list of vectors cotaining $\mathbf{0}$ is linearly dependent.

3. Since $x \sim \mathcal{N}(2, 1)$ independent of $y \sim \mathcal{N}(1, 1)$, together they are multivariate normal:

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right),$$

where $\mathrm{Cov}(x, y) = \mathrm{Cov}(y, x) = 0$ since $x \perp y$.

Put $\boldsymbol{\mu} = [2, 1]'$ so that we can write $[x, y]' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{I}_{2\times 2})$. Using the results on non-central chi-square distributions (slide 35 of set 1),

$$x^2 + y^2 = \left[x, y\right] \begin{bmatrix} x \\ y \end{bmatrix} \sim \chi_2^2(\boldsymbol{\mu}'\boldsymbol{\mu}/2),$$

where the non-centrality parameter reduces to

$$\boldsymbol{\mu}'\boldsymbol{\mu}/2 = \frac{1}{2}\left[2, 1\right] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{5}{2}.$$

This gives us $x^2 + y^2 \sim \chi_2^2(2.5)$, so that

$$\mathrm{P}\left(\sqrt{x^2 + y^2} > 6\right) = \mathrm{P}\left(x^2 + y^2 > 36\right) = \mathrm{P}\left(\chi_2^2(2.5) > 36\right).$$

We can compute the above probability in `R` using the `pchisq()` function. The documentation brought up by `?pchisq` says

> "The non-central chi-squared distribution with df $= n$ degrees of freedom and non-centrality parameter ncp $= \lambda$ [...] this is the distribution of the sum of squares of n normals each with variance one, λ being the sum of squares of the normal means."

This implies `R` parameterizes the non-centrality parameter using $\boldsymbol{\mu}'\boldsymbol{\mu} = \mu_1^2 + \cdots + \mu_n^2$ (this is the sum of squares of normal means) rather than $\boldsymbol{\mu}'\boldsymbol{\mu}/2$. So, we need to double the non-centrality parameter we obtained above when inputing the function arguments. The code

```
> pchisq(36, df = 2, ncp = 5, lower.tail = FALSE)
```

gives the result

$$\mathrm{P}\left(\sqrt{x^2 + y^2} > 6\right) = 0.00014096.$$

Comments: Some students did a Monte Carlo experiment to approximate this probability; others did a change of variable transformation or used a moment generating function to establish the distribution of x^2+y^2. Noticing that we can write x^2+y^2 in terms of a multivariate normal vector and applying the results in the slide set, as done above, is not an approximation (as Monte Carlo methods are), and may be easier than a change of variable transformation or using a moment generating function.

4. We have $z_1, z_2 \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. By slide 31 of set 1,

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim \mathcal{N}(\mathbf{0}_{2\times 2}, \boldsymbol{I}_{2\times 2}).$$

   (a) Let $\boldsymbol{A} = \left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right]$. By slide 32 of set 1,

$$\frac{1}{\sqrt{2}}(z_1 - z_2) = \frac{1}{\sqrt{2}}\,[1, -1]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \left[\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}\right]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \boldsymbol{A}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim \mathcal{N}(0, \boldsymbol{A}\boldsymbol{A'}).$$

   Since $\boldsymbol{A}\boldsymbol{A'} = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \left(-\frac{1}{\sqrt{2}}\right) \cdot \left(-\frac{1}{\sqrt{2}}\right) = 1$, we have

$$\frac{1}{\sqrt{2}}(z_1 - z_2) \sim \mathcal{N}(0,1).$$

   Squaring, by slide 36 of set 1, we obtain a central chi-square random variable:

$$\frac{1}{2}(z_1 - z_2)^2 \sim \chi^2_1.$$

   (b) Notice that

$$\frac{z_1 + z_2}{|z_1 - z_2|} = \frac{\frac{1}{\sqrt{2}}(z_1 + z_2)}{\sqrt{\frac{1}{2}(z_1 - z_2)^2/1}}. \qquad (1)$$

   Similar to the first step in part (a), the numerator in (1) is standard normal:

$$\frac{1}{\sqrt{2}}(z_1 + z_2) \sim \mathcal{N}(0,1).$$

   From the result in part (a), the part of the denominator in (1) under the square root is a central chi-square on one degree of freedom.

   To have a t_1 distribution, we need to show that the random variables in the numerator and under the square root in the denominator in (1) are independent. We can do this by using slide 44 of slide set 1. Let $\boldsymbol{z} = [z_1, z_2]'$. Then $\frac{1}{\sqrt{2}}(z_1 + z_2) = \boldsymbol{A}_1\boldsymbol{z}$ where $\boldsymbol{A}_1 = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]$. Also,

$$\frac{1}{2}(z_1 - z_2)^2 = \frac{1}{2}\begin{bmatrix} z_1 & z_2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \boldsymbol{z}'\boldsymbol{A}_2\boldsymbol{z}$$

   where

$$\boldsymbol{A}_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Now note that $\boldsymbol{A}_1\boldsymbol{I}_{2\times 2}\boldsymbol{A}_2 = \mathbf{0}_{1\times 2}$. Consequently, $\frac{1}{\sqrt{2}}(z_1+z_2)$ is independent of $\frac{1}{2}(z_1-z_2)^2$ by slide 44 of slide set 1. By the result on slide 40 of set 1,

$$\frac{z_1+z_2}{|z_1-z_2|} = \frac{\frac{1}{\sqrt{2}}(z_1+z_2)}{\sqrt{\frac{1}{2}(z_1-z_2)^2/1}} \sim t_1.$$

Comments:

- Many students failed to check the independence condition needed to establish a t_1 distribution in (b).
- Some students did a change of variable transformation or used a moment generating function. Others showed that $\frac{z_1+z_2}{|z_1-z_2|}$ is a ratio of independent standard normal random variables and hence Cauchy(0,1), which is the same distribution as t_1.
- Another way to show that the random variables in the numerator and under the square root in the denominator in (1) are independent:
  Since
  $$\begin{bmatrix} z_1+z_2 \\ z_1-z_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \right),$$
  $Cov(z_1+z_2, z_1-z_2) = 0$ implies that z_1+z_2 and z_1-z_2 are independent. Thus, $\frac{1}{\sqrt{2}}(z_1+z_2)$ and $\sqrt{\frac{1}{2}(z_1-z_2)^2}$ are independent.

5. We are given $y_1, \dots, y_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$.

(a) Using properties of matrix transposition,

$$
\begin{aligned}
s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y}_\cdot)^2 \\
&= \frac{1}{n-1}(\boldsymbol{y} - \bar{y}_\cdot \mathbf{1}_{n\times 1})'(\boldsymbol{y} - \bar{y}_\cdot \mathbf{1}_{n\times 1}) \\
&= \frac{1}{n-1}\left(\boldsymbol{y} - \frac{1}{n}\sum_{i=1}^{n} y_i \mathbf{1}_{n\times 1}\right)'\left(\boldsymbol{y} - \frac{1}{n}\sum_{i=1}^{n} y_i \mathbf{1}_{n\times 1}\right) \\
&= \frac{1}{n-1}\left(\boldsymbol{y} - \frac{1}{n}\mathbf{1}_{n\times 1}\sum_{i=1}^{n} y_i\right)'\left(\boldsymbol{y} - \frac{1}{n}\mathbf{1}_{n\times 1}\sum_{i=1}^{n} y_i\right) \\
&= \frac{1}{n-1}\left(\boldsymbol{y} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\boldsymbol{y}\right)'\left(\boldsymbol{y} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\boldsymbol{y}\right) \\
&= \frac{1}{n-1}\left[\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right)\boldsymbol{y}\right]'\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right)\boldsymbol{y} \\
&= \frac{1}{n-1}\boldsymbol{y}'\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right)'\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right)\boldsymbol{y} \\
&= \boldsymbol{y}'\left[\frac{1}{n-1}\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right)'\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right)\right]\boldsymbol{y} \\
&= \boldsymbol{y}'\boldsymbol{B}\boldsymbol{y},
\end{aligned}
$$

where

$$
\boldsymbol{B} = \frac{1}{n-1}\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right)'\left(\boldsymbol{I}_{n\times n} - \frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}'\right).
$$

Notice that $\boldsymbol{B} = \frac{1}{n-1}\left(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\mathbf{1}_{n\times 1}}\right)$, because

$$
\begin{aligned}
\frac{1}{n}\mathbf{1}_{n\times 1}\mathbf{1}_{n\times 1}' &= \mathbf{1}_{n\times 1}(n^{-1})\mathbf{1}_{n\times 1}' \\
&= \mathbf{1}_{n\times 1}(\mathbf{1}_{n\times 1}'\mathbf{1}_{n\times 1})^{-1}\mathbf{1}_{n\times 1}' \\
&= \boldsymbol{P}_{\mathbf{1}_{n\times 1}}.
\end{aligned}
$$

Comments: As seen above, we do not need to assume any model for $\boldsymbol{y}$ (e.g., we do not need the GMMNE on slide 16 of set 2) to show the existence of $\boldsymbol{B}$ such that $s^2 = \boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}$.

(b) From part (a), we have $\boldsymbol{B} = \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\mathbf{1}_{n\times 1}})$. Since $y_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, it follows that

$$
[y_1, \dots, y_n]' = \boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right),
$$

where $\boldsymbol{\mu} = \mu\mathbf{1}_{n\times 1}$ and $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_{n\times n}$. Clearly $\boldsymbol{\Sigma}$ is positive definite, assuming that $\sigma^2 > 0$:

for any non-zero $\boldsymbol{t} = [t_1, \ldots, t_n]' \in \mathbb{R}^n$ (i.e., $\boldsymbol{t} \neq \boldsymbol{0}_{n\times 1}$), we have

$$\begin{aligned}
\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t} &= \boldsymbol{t}'[\sigma^2\boldsymbol{I}_{n\times n}]\boldsymbol{t} \\
&= \sigma^2\boldsymbol{t}'\boldsymbol{I}_{n\times n}\boldsymbol{t} \\
&= \sigma^2\boldsymbol{t}'\boldsymbol{t} \\
&= \sigma^2(t_1^2 + \cdots + t_n^2) \\
&> 0
\end{aligned}$$

since $\boldsymbol{t} \neq \boldsymbol{0}_{n\times 1}$ implies $t_i \neq 0$ for at least one $i \in \{1, \ldots, n\}$.

Set $\boldsymbol{A} = \frac{n-1}{\sigma^2}\boldsymbol{B}$. By properties of rank and the results on slide 5 of set 2 that pertain to the rank of an orthogonal projection matrix,

$$\begin{aligned}
\operatorname{rank}(\boldsymbol{A}) &= \operatorname{rank}\left(\frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\right) \\
&= \operatorname{rank}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= \operatorname{tr}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= \operatorname{tr}(\boldsymbol{I}_{n\times n}) - \operatorname{tr}(\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= n - \operatorname{rank}(\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= n - \operatorname{rank}(\boldsymbol{1}_{n\times 1}) \\
&= n - 1.
\end{aligned}$$

As an orthogonal projection matrix, $\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}$ is symmetric, and clearly $\boldsymbol{I}_{n\times n}$ is symmetric. Hence $\boldsymbol{A}$ is also symmetric. We also have that $\boldsymbol{A\Sigma}$ is idempotent:

$$\begin{aligned}
\boldsymbol{A\Sigma A\Sigma} &= \frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\left[\sigma^2\boldsymbol{I}_{n\times n}\right]\frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\left[\sigma^2\boldsymbol{I}_{n\times n}\right] \\
&= (\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= (\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= \frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\left[\sigma^2\boldsymbol{I}_{n\times n}\right] \\
&= \boldsymbol{A\Sigma},
\end{aligned}$$

since $\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}$ is idempotent (this is easy to show using the fact that $\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}$ is idempotent).

We have now established all the ingredients that we need to apply the result on slide 31 of set 1. Hence,

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \frac{n-1}{\sigma^2}s^2 \sim \chi^2_{n-1}(\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2).$$

The non-centrality parameter reduces to zero:

$$
\begin{aligned}
\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2 &= [\mu\mathbf{1}_{n\times 1}]' \frac{n-1}{\sigma^2}\frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\mathbf{1}_{n\times 1}})\,[\mu\mathbf{1}_{n\times 1}]\cdot\frac{1}{2} \\
&= \frac{\mu^2}{2\sigma^2}\mathbf{1}'_{n\times 1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\mathbf{1}_{n\times 1}})\mathbf{1}_{n\times 1} \\
&= \frac{\mu^2}{2\sigma^2}(\mathbf{1}'_{n\times 1}\boldsymbol{I}_{n\times n}\mathbf{1}_{n\times 1} - \mathbf{1}'_{n\times 1}\boldsymbol{P}_{\mathbf{1}_{n\times 1}}\mathbf{1}_{n\times 1}) \\
&= \frac{\mu^2}{2\sigma^2}(\mathbf{1}'_{n\times 1}\mathbf{1}_{n\times 1} - \mathbf{1}'_{n\times 1}\left[\mathbf{1}_{n\times 1}(\mathbf{1}'_{n\times 1}\mathbf{1}_{n\times 1})^{-1}\mathbf{1}'_{n\times 1}\right]\mathbf{1}_{n\times 1}) \\
&= \frac{\mu^2}{2\sigma^2}(\mathbf{1}'_{n\times 1}\mathbf{1}_{n\times 1} - \mathbf{1}'_{n\times 1}\mathbf{1}_{n\times 1}) \\
&= \frac{\mu^2}{2\sigma^2}(n-n) \\
&= 0,
\end{aligned}
$$

proving the desired result that

$$\frac{n-1}{\sigma^2}s^2 \sim \chi^2_{n-1}.$$

Comments: Many students didn't establish all the conditions listed on slide 35, such as $\boldsymbol{A}$ is symmetric or that $\boldsymbol{y}$ is multivariate normal. A few students assumed the non-centrality parameter was zero without any mention of $\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2$, and others only showed minimal work in establishing that $\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2 = 0$.

6. Consider a matrix

$$\boldsymbol{A}_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

so that

$$
\begin{aligned}
\boldsymbol{A}'\boldsymbol{A} &= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^m a_{i1}^2 & \sum_{i=1}^m a_{i1}a_{i2} & \cdots & \sum_{i=1}^m a_{i1}a_{in} \\ \sum_{i=1}^m a_{i1}a_{i2} & \sum_{i=1}^m a_{i2}^2 & \cdots & \sum_{i=1}^m a_{i2}a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{i1}a_{in} & \sum_{i=1}^m a_{i2}a_{in} & \cdots & \sum_{i=1}^m a_{in}^2 \end{bmatrix}.
\end{aligned}
$$

$\Rightarrow$: ["only if" part] Suppose that $\boldsymbol{A} = \mathbf{0}_{m\times n}$. Clearly then $\boldsymbol{A}'_{n\times m} = \mathbf{0}_{n\times m}$, which implies $\boldsymbol{A}'\boldsymbol{A} = \mathbf{0}_{n\times m}\mathbf{0}_{m\times n} = \mathbf{0}_{n\times n}$.

$\Leftarrow$: ["if" part] Suppose that $\boldsymbol{A}'\boldsymbol{A} = \mathbf{0}_{n\times n}$. This requires that the diagonal elements of $\boldsymbol{A}'\boldsymbol{A}$ are only zeros, that is, $\sum_{i=1}^m a_{ij}^2 = 0$ for $j = 1, \ldots, n$. Consequently, $a_{ij} = 0$ for

$j = 1, \ldots, n,\ i = 1, \ldots, m$, which implies $\boldsymbol{A} = \mathbf{0}_{m \times n}$.

Thus, $\boldsymbol{A} = \mathbf{0} \iff \boldsymbol{A}'\boldsymbol{A} = \mathbf{0}$.

Comments: This question, as well as the next one, requires a proof of an "if and only if" statement. This means you need to consider both directions (both the "if" and "only if" parts) to prove the desired conclusion. While it is trivial to many that $\boldsymbol{A} = \mathbf{0} \implies \boldsymbol{A}'\boldsymbol{A} = \mathbf{0}$, if you don't include this in your proof, you can't claim to have proven the desired result.

7. $\Leftarrow$: ["if" part] Suppose $\boldsymbol{XA} = \boldsymbol{XB}$. This part of the proof is trivial: multiplying both sides by $\boldsymbol{X}'$ gives $\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB}$.

   $\Rightarrow$: ["only if" part] Conversely, suppose $\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB}$. By properties of matrix algebra and transpose, as well as the result of part (a), we have

$$
\begin{aligned}
\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB} &\implies \boldsymbol{X}'\boldsymbol{XA} - \boldsymbol{X}'\boldsymbol{XB} = \mathbf{0} \\
&\implies \boldsymbol{X}'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \mathbf{0} \\
&\implies (\boldsymbol{A} - \boldsymbol{B})'\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \mathbf{0} \\
&\implies \big(\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B})\big)'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \mathbf{0} \\
&\implies \boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \mathbf{0} && \text{by part (a)} \\
&\implies \boldsymbol{XA} - \boldsymbol{XB} = \mathbf{0} \\
&\implies \boldsymbol{XA} = \boldsymbol{XB}.
\end{aligned}
$$

   Therefore, $\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB} \iff \boldsymbol{XA} = \boldsymbol{XB}$.

   Comments: As in problem 6, a proof of an"if and only if" statement requires both directions to be complete. Additionally, note that depending on the matrix dimensions, we may have

$$
\big(\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B})\big)'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \mathbf{0}_{n \times n} \neq \mathbf{0}_{m \times n} = \boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}),
$$

   where $m, n \in \{1, 2, \ldots\}$.