Exam 02409 Multivariate Statistics - E21

6th of December, 2021
02409 - Multivariate Statistics

4 hours written exam - All aids allowed, including internet access.

There are 6 problems with a total of 30 questions.
The questions in a specific problem should be read in order.

The questions are weighted equally

A correct answer gives 5 points, a wrong answer gives –1 point. Unanswered questions or a “don’t know” give 0 points. The total number of points needed for a satisfactorily answered exam is determined at the final evaluation of the exam. Especially note that the grade 10 may be given even if only one answer is wrong or unanswered.

For each problem there is a link to an enclosure or a dataset, if needed to solve that problem.

PROBLEM 1

[Enclosure A (https://resources.mcq.eksamen.dtu.dk/v1/82169650-6fdc-4c1b-a8a3-f5e7f4a7a46b)](https://resources.mcq.eksamen.dtu.dk/v1/82169650-6fdc-4c1b-a8a3-f5e7f4a7a46b) with SAS program and S/
belongs to this problem.

We consider the relation between education and voting pattern. For each of the 98 Danish municipalities, we have the highest level of education each person has and how many votes each party got, at the municipal election in 2(data is from [Danmark Statistik (https://www.statistikbanken.dk)](https://www.statistikbanken.dk).

The following variables are included in our analysis. It is *not* important to understand the meaning of the variables

Educational variables	Meaning
H10	9 years mandatory school
H20	High School
H30	Vocational School
H35	Admittance education for universities
H40	Short higher education
H50	Medium higher education
H60	Bachelor
H70	Long higher education
H80	PhD
H90	Unknown

Political variables	Meaning
A	Socialdemokraterne
B	Radikale Venstre
C	Konservative
D	Nye Borgerlige
F	Socialistisk Folkeparti
I	Liberal Alliance
O	Dansk Folkeparti
V	Venstre
Ø	Enhedslisten
Å	Alternativet

We perform a canonical correlation analysis.

Question 1.1

The 3rd set of canonical variates explain the following fraction of variation between the two datasets

Choose one answer

- 0.9969
- Don't know
- 0.06363783
- 0.857140
- 0.734690
- 0.0045

Question 1.2 - Enclosure A

Using the usual notation, the following number of eigenvalues of $\mathbf{E}^{-1}\mathbf{H}$ are significantly different from zero at the 5%-level:

Choose one answer

- 2
- Don't know
- 4
- 5
- 1
- 3

Question 1.3 - Enclosure A

What fraction of variance in the WITH variable 'C' is explained by V1 and V2 together

Choose one answer

- 19 %
- 98 %
- 11 %
- 100 %
- 26 %
- Don't know

Question 1.4 - Enclosure A

The interpretation - when using the appropriate correlations - of V4 vs W4 is broadly

Choose one answer

- V4 is rougly an average of all VAR variables, and W4 is a contrast between (Ø,Å) vs. A
- Don't know
- V4 is mainly a contrast between shorter education (H10, H30, H35, H40, H50) vs. long (H60, H70, H80), and W4 is a contrast between B vs. I
- V4 is a contrast between H10 vs H20, and W4 is an average of all WITH variables
- V4 is mainly negatively correlated with admittance education (H35), and W4 is a contrast between C vs. (A, F)
- V4 is mainly a contrast between shorter education (H10, H30, H35, H40, H50) vs. long (H60, H70, H80), and W4 is a contrast between A vs. B

PROBLEM 2

You are encouraged to use statistical software to answer this problem.

We still consider the data from Problem 1, but now only a small subset, namely the 17 municipalities in RegionMid
data can be found here (https://resources.mcq.eksamen.dtu.dk/v1/5b608694-45c5-413d-b57d-08693c275379) .

We will now try to predict the election results for the party A (socialdemokraterne) based on the educational level
municipalities.

Question 2.1
We now perform a linear regression with A as the dependent variable and including all educational levels depende
variables as well as an intercept, i.e., a model of the form

$$A = \mu + \beta_1 H10 + \beta_2 H20 + \beta_3 H30 + \beta_4 H35 + \beta_5 H40 + \beta_6 H50 + \beta_7 H60 + \beta_8 H70 + \beta_9 H80 + \beta_{10} H$$

Where μ is the intercept and ϵ is the error term.

The variable with the largest tolerance in this model is:

Choose one answer

- H90
- H80
- H40
- H30
- H20
- Don't know

Question 2.2

We still consider the data from Problem 2

When only including four dependent variables, we attain the highest R^2 by including

Choose one answer

- H10 H35 H40 H60
- H35 H60 H80 H90
- H10 H20 H35 H40
- Don't know
- H20 H40 H60 H70
- H10 H30 H60 H90

Question 2.3

We still consider the data from Problem 2

We now perform a linear regression with A as the dependent variable and H20 H60 H80 as the independent variables with an intercept, i.e., a model of the form:
$A = \mu + \gamma_1 H20 + \gamma_2 H60 + \gamma_3 H80 + \epsilon$
Where μ is the intercept and ϵ is the error term.

The observation with the *lowest* leverage is:

Choose one answer

- Don't know
- 16
- 11
- 10
- 7
- 1

Question 2.4

We still consider the data from Problem 2

We still perform a linear regression with A as the dependent variable and H20 H60 H80 as the independent variables with an intercept, i.e., a model of the form:
$A = \mu + \gamma_1 H20 + \gamma_2 H60 + \gamma_3 H80 + \epsilon$
Where μ is the intercept and ϵ is the error term.

The observation with largest impact on the parameter estimate for H80 is:

Choose one answer

- 1
- 7
- 11
- Don't know
- 10
- 16

Question 2.5

We still consider the data from Problem 2

We still perform a linear regression with A as the dependent variable and H20 H60 H80 as the independent variables with an intercept, i.e., a model of the form:
$A = \mu + \gamma_1 H20 + \gamma_2 H60 + \gamma_3 H80 + \epsilon$
Where μ is the intercept and ϵ is the error term.

The number of observations with a Cook's D larger than 0.235 is:

Choose one answer

- Don't know
- 3
- 5
- 2
- 4
- 1

Question 2.6

We still consider the data from Problem 2

We still perform a linear regression with A as the dependent variable and H20 H60 H80 as the independent variables with an intercept, i.e., a model of the form:
$A = \mu + \gamma_1 H20 + \gamma_2 H60 + \gamma_3 H80 + \epsilon$
Where μ is the intercept and ϵ is the error term.

The two parameter estimates with the numerically largest covariance are:

Choose one answer

- H20 and H80
- H20 and H60
- Intercept and H80
- Don't know
- Intercept and H60
- H60 and H80

PROBLEM 3

Enclosure B (https://resources.mcq.eksamen.dtu.dk/v1/a7baa2f6-8563-4555-8485-0788a5b6e1e6) with SAS program and SAS output belongs to this problem.

We now look at the relation between different glass types and the mineral content of the glass and the refractive index. The data is from https://archive.ics.uci.edu/ml/datasets/glass+identification (https://archive.ics.uci.edu/ml/datasets/glass+identification)
and was collected by *B. German, Central Research Establishment, Home Office Forensic Science Service, Aldermaston, Reading, Berkshire RG7 4PN*

We consider the following types and mineral contents. A detailed understanding of the types of glass and mineral content is not necessary.

The six (6) different types of glass are (note ID 4 is missing)

Glass ID	Type
1	Building window, float processed
2	Building window, non float processed
3	Vehicle window, float processed
5	Containers
6	Table ware
7	Headlamps

The classification variables are

Classification variable	Meaning
RI	Refractive Index
Na	Sodium
Mg	Magnesium
Al	Aluminium
Si	Silicon
K	Potassium
Ca	Calcium
Ba	Barium
Fe	Iron

We will try to distinguish between the different glass types by means of discriminant analyses.

Question 3.1

The number of resubstitution misclassifications - when going from Quadratic Discriminant Analysis to Linear Discriminant Analysis - increases by:

Choose one answer

- 21
- Don't know
- 13
- 0
- 6
- 37

Question 3.2 - Enclosure B

The two most different classes as measured by Mahalanobis distances (based on the pooled dispersion matrix) are:

Choose one answer

- ○ 1 and 7
- ○ 3 and 7
- ○ 1 and 3
- ○ 5 and 6
- ○ Don't know
- ○ 2 and 3

Question 3.3 - Enclosure B

Regardless of the previous question we now test if type 1 and type 3 have different mean values, when using all nine classification variables.

The p-value for the usual test of difference in means is:

Choose one answer

- 0.0023
- 0.0043
- 0.1142
- < 0.0001
- $9 \cdot \frac{2.04821}{70+17}$

Question 3.4 - Enclosure B

Considering the Linear Discriminant Analysis, we now test whether Ca, Fe, and Si contribute to the discrimination between type 1 and type 3. The usual test statistic for this is:

Choose one answer

- $\frac{70+17-9-1}{9-3} \cdot \frac{2.16348-1.11352}{(70+17)(70+17-2)/(70\cdot 17)+1.11352}$
- $1.23907 - 0.85839$
- $\frac{70+17-9-1}{9-6} \cdot \frac{2.16348-1.11352}{(70+17)(70+17-2)/(70\cdot 17)+1.11352}$
- $\frac{70+76-9-1}{9-6} \cdot \frac{2.16348-0.17698}{(70+76)(70+76-2)/(70\cdot 76)+0.17698}$
- Don't know
- 3.0076

Question 3.5 - Enclosure B

We now consider the Quadratic Discriminant Analysis. When going from the LDA to the QDA (both including all variables and classes) the resubstitution specificity of class 3 is increased by:

Choose one answer

- $\frac{13}{214}$
- Don't know
- 0.8086
- 0.0318
- $\frac{16-11}{23-10}$
- $\frac{16-11}{214}$

PROBLEM 4

Enclosure C (https://resources.mcq.eksamen.dtu.dk/v1/b46a4076-cc95-4a54-95f0-c1f3ac23857c) with SAS program and SAS output belongs to this problem.

We still consider the glass data introduced in Question 3. However, we now only consider the five types 1, 2, 3, 5, and 7.
Further, we have also omitted the refractive index (RI) from the analysis.

We will now investigate the relationship between the mineral contents.

Question 4.1

The usual test statistic for the correlation between Mg and Ba conditioned on Al being different from zero is:

Choose one answer

- -0.3585
- Don't know
- -5.4580
- -0.3840
- 14.2127
- -5.4715

Question 4.2 - Enclosure C

When performing a PCA on the correlation matrix of the data, we need the following number of components to account for at least 90 % of the variance in the data

Choose one answer

- 1
- 7
- 2
- Don't know
- 5
- 4

Question 4.3 - Enclosure C

We still consider a PCA on the correlation matrix of the data.

The usual test-statistic for the last two eigenvalues being equal is:

Choose one answer

- $\frac{0.30794515-0.00172330}{205}$
- 781.16
- 0.0036
- 1123.8
- 1163.6
- Don't know

Question 4.4 - Enclosure C

We now consider a factor analysis on the data.

Unrotated factor 1 and 2 combined, explains the following fraction of the variance in *Al*

Choose one answer

- $\frac{2.2703185+1.5881626}{8}$
- Don't know
- 0.6551
- 0.8212
- 0.49715^2
- 0.6954

Question 4.5 - Enclosure C

We still consider a factor analysis.

The interpretation of the rotated factors is:

Choose one answer

- F1 is an average of all variables except Mg
  F2 is mainly a contrast between Ca, Fe, K and Si
  F3 is mainly a contrast between K and Ca
  F4 is a contrast between Ca vs. Si
- F1 is mainly a contrast between Al and Ba vs. Mg
  F2 is mainly a contrast between Ca and Fe vs. Na
  F3 is mainly a contrast between K and Ca
  F4 is mainly Si
- Don't know
- F1 is mainly a contrast between Al and Ba vs. Mg
  F2 is an average of Ba, K, Al, and Mg
  F3 is an average of Ca and K
  F4 is mainly Si
- F1 is mainly a contrast between Ca and Fe vs. Na
  F2 is mainly a contrast between Al and Ba vs. Mg
  F3 is mainly a contrast between K and Ca
  F4 is mainly Si
- F1 is mainly a contrast between Mg vs. all other variables
  F2 is mainly Na
  F3 is an average of Na, Ba, Fe, and Ca
  F4 is a contrast between K vs. Si

PROBLEM 5

Enclosure D (https://resources.mcq.eksamen.dtu.dk/v1/05dc6207-23f7-4510-aacc-558de016b2c5) with SAS program and SAS output belongs to this problem.

We now consider high-school students and the connection between 1) the parents' educational level and 2) free vs. payed lunch on the grades in math, reading and writing.

The data are from https://www.kaggle.com/roshansharma/student-performance-analysis (https://www.kaggle.com/roshansharma/student-performance-analysis)

We only consider the first 100 observations. It is not important to understand the variables in detail.

The parental levels are

Name	Meaning - Parents highest level of education is an:
associat	associate degree
bachelor	bachelor degree
college	college degree
highscho	high school degree
master	masters degree

We consider a model of the form:

$$[\text{math reading writing}] = \mu + \text{parents}_k + \text{lunch}_m$$

Question 5.1

We now consider the parental effect on grades. The p-value for the usual test for parental effect falls in the interval:

Choose one answer

-]0.01 - 0.05]
-]0.001 - 0.01]
- Don't know
-]0.1 - 0.5]
-]0.0001 - 0.001]
-]0.05 - 0.1]

Question 5.2 - Enclosure D

Using the usual notation, the matrix Q1 is given by

Choose one answer

- ○ $\begin{bmatrix} 0.00307 & -0.01789 & 0.01863 \\ -0.01288 & 0.00852 & 0.00550 \\ 0.00565 & 0.01579 & -0.01538 \end{bmatrix}$
- ○ $\begin{bmatrix} 3074.7 & 2903.7 & 2722.3 \\ 2903.7 & 2742.2 & 2570.8 \\ 2722.3 & 2570.8 & 2410.2 \end{bmatrix}$
- ○ $\begin{bmatrix} 20364 & 17783 & 18668 \\ 17783 & 20912 & 20942 \\ 18668 & 20942 & 22743 \end{bmatrix}$
- ○ $\begin{bmatrix} 0.00631 & 0.00837 & -0.00765 \\ -0.01282 & 0.01276 & 0.00087 \\ -0.00175 & -0.02022 & 0.02355 \end{bmatrix}$
- ○ $\begin{bmatrix} 1351.7 & 1044.9 & 1530.6 \\ 1044.9 & 863.54 & 1231.8 \\ 1530.6 & 1231.8 & 1796.1 \end{bmatrix}$
- ○ Don't know

Question 5.3 - Enclosure D

We now consider the lunch effect. The usual test statistic for lunch effect, if we only consider the variables *math* and *reading,* is:

Choose one answer

- ○ 0.7567
- ○ 14.19+12.33
- ○ $\frac{4949.77042+3946.95115}{20363.53958+20911.68885}$
- ○ Don't know
- ○ 0.9259
- ○ 0.8666

PROBLEM 6

We consider a normally distributed random variable

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{bmatrix}$$

with expectation vector and dispersion matrix equal to

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

In the sequel, you may find the following expressions useful

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} - \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 10 & 8 \\ 8 & 26 \end{bmatrix}$$

Question 6.1

The squared correlation between Y_1 and Y_2 is

Choose one answer

- ○ Don't know
- ○ $\frac{1}{8}$
- ○ $\frac{1}{2}$
- ○ $\frac{1}{16}$
- ○ $\frac{1}{64}$
- ○ $\frac{1}{4}$

Question 6.2

We still consider a normally distributed random variable

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{bmatrix}$$

with expectation vector and dispersion matrix equal to

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

In the sequel, you may find the following expressions useful

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$
$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} - \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 10 & 8 \\ 8 & 26 \end{bmatrix}$$

The conditional mean $E(Y_1|X_2 = x_2)$ is

Choose one answer

- x_1
- Don't know
- x_2
- 0
- $\frac{5}{7}x_1 - \frac{13}{7}x_2$
- $\frac{1}{7}x_1 + \frac{2}{7}x_2$

Question 6.3

We still consider a normally distributed random variable

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{bmatrix}$$

with expectation vector and dispersion matrix equal to

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

In the sequel, you may find the following expressions useful

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} - \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 10 & 8 \\ 8 & 26 \end{bmatrix}$$

The conditional mean $E(Y_1|X_1 = x_1)$ is

Choose one answer

- ○ $4x_1 - x_2 - 3$
- ○ Don't know
- ○ $\frac{4}{7}x_1 - \frac{1}{7}x_2$
- ○ $\frac{1}{2}x_1 - \frac{1}{2}$
- ○ $\frac{1}{2}x_1$
- ○ 0

Question 6.4

We still consider a normally distributed random variable

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{bmatrix}$$

with expectation vector and dispersion matrix equal to

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

In the sequel, you may find the following expressions useful

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$
$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} - \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 10 & 8 \\ 8 & 26 \end{bmatrix}$$

The conditional mean $E(Y_1|\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix})$ is

Choose one answer

- Don't know
- $\frac{4}{7}x_1 - \frac{1}{7}x_2 - \frac{3}{7}$
- 0
- x_1
- x_2
- $4x_1 - 2x_2$

Question 6.5

We still consider a normally distributed random variable

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{bmatrix}$$

with expectation vector and dispersion matrix equal to

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

In the sequel, you may find the following expressions useful

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$
$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} - \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 10 & 8 \\ 8 & 26 \end{bmatrix}$$

The squared partial correlation $\rho^2_{Y_1 Y_2 | X_1}$ between Y_1 and Y_2 given X_1 is

Choose one answer

- $\frac{2}{7}$
- $\frac{64}{260}$
- Don't know
- $\frac{1}{2}$
- $\frac{1}{6}$
- $\frac{260}{\sqrt{(1-(7/8)^2)(1-(7/8)^2)}}$

Question 6.6

We still consider a normally distributed random variable

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{bmatrix}$$

with expectation vector and dispersion matrix equal to

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

In the sequel, you may find the following expressions useful

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$
$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} - \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 10 & 8 \\ 8 & 26 \end{bmatrix}$$

The squared partial correlation $\rho^2_{Y_1 Y_2 | X_1 X_2}$ between Y_1 and Y_2 given **X** is

Choose one answer

- $\frac{64}{260}$
- $\frac{8}{49}$
- Don't know
- $\frac{1}{\sqrt{2 \cdot 4}}$
- $\frac{5}{7}$
- $\frac{260}{\sqrt{(1-(7/8)^2)(1-(7/8)^2)}}$

Question 6.7

We still consider a normally distributed random variable

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{bmatrix}$$

with expectation vector and dispersion matrix equal to

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

In the sequel, you may find the following expressions useful

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}^{-1} = \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} - \frac{1}{7}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 10 & 8 \\ 8 & 26 \end{bmatrix}$$

The squared multiple correlation $\rho^2_{Y_1|X_1X_2}$ between Y_1 and $[X_1 \ X_2]^T$ is

Choose one answer

- $\frac{1 \cdot 96 \cdot 1}{16 \cdot 7 \cdot 1}$
- $\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} / 7$
- Don't know
- $\frac{2}{7}$
- $\frac{2(1-(7/8)^2)}{4}$
- $\frac{7}{260}$