

UNIVERSITY OF BRISTOL

May/June 2020 Examination Period

FACULTY OF ENGINEERING

Second Year Examination for the Degrees of BSc, BEng and MEng

COMS21202

Symbols, Patterns and Signal

TIME ALLOWED:
2 hours

This paper contains 20 multiple choice questions worth a maximum of 60 marks
The number of marks allocated for each multiple choice question is indicated

For each multiple choice, choose only one answer

Each question has exactly one correct answer. You must select exactly one answer per question, anything else will not be counted. You must mark your answers with a pencil and only as indicated on the image below.

**All questions will be used for assessment.
Calculators must have the Engineering Faculty seal of approval.**

INSTRUCTIONS		CANDIDATE NO.	CANDIDATE NUMBER	CANDIDATE NO.
This form is designed to be machine readable.			Your candidate number is a 5-digit number.	0 2 4 6 8
Please mark like this — and use only an HB pencil.			Complete the grid with your candidate number.	— 0 0 0 0 0
If you make a mistake use an eraser.			For example, a student with the candidate number 02468 would fill the box as illustrated.	0 2 4 6 8
Do not mark the paper outside the boxes provided.			0 2 4 6 8	0 2 4 6 8

TURN OVER ONLY WHEN TOLD TO START WRITING

Help Formulas:

One-dimensional Gaussian/Normal probability density function:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Multi-dimensional Gaussian/Normal probability density function:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^M |\Sigma|}} e^{-\frac{1}{2} (\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}$$

Least Squares Matrix Form:

$$\mathbf{a}_{LS} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

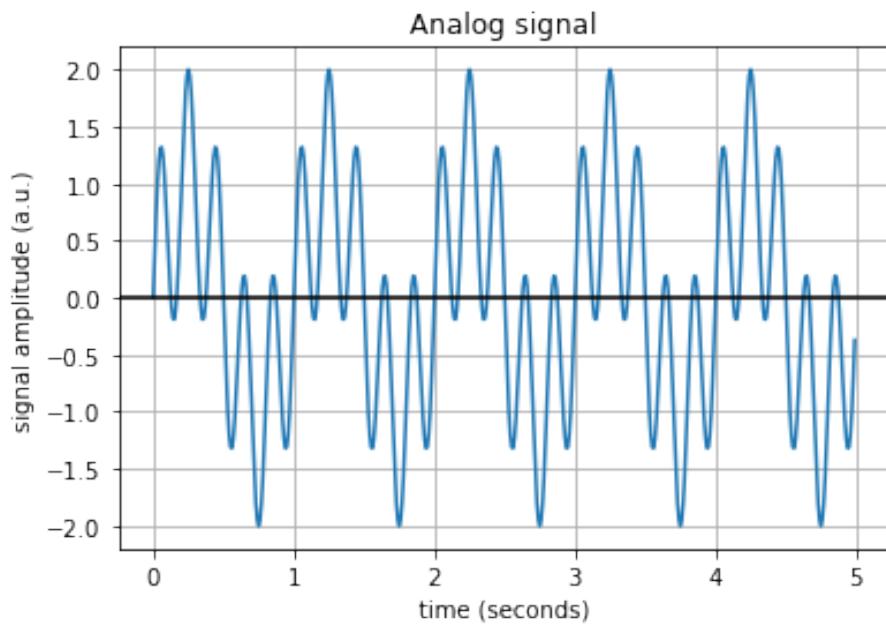
Matrix inversion:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Matrix Determinant:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

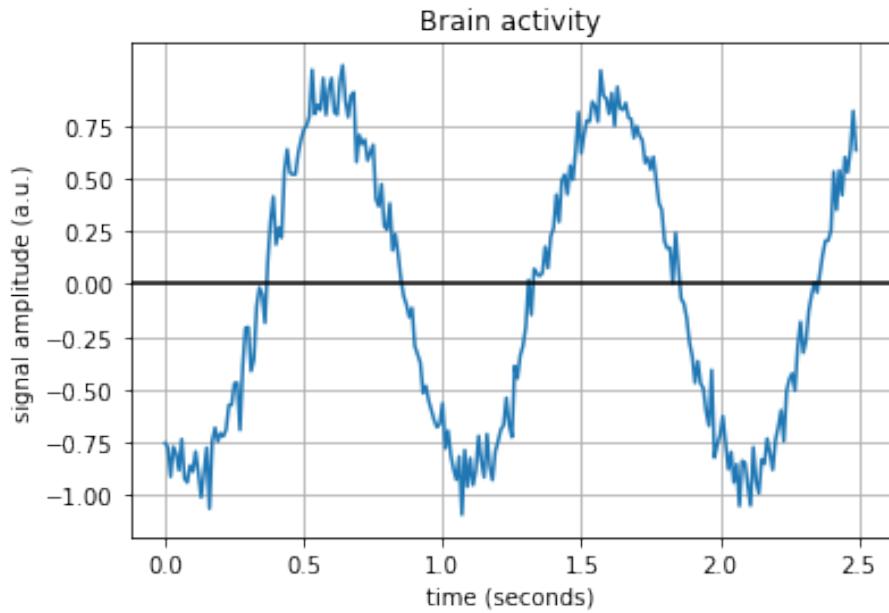
1. The image below shows an analogue one-dimensional signal. You are asked to digitise this signal. For that you have to decide which sampling frequency and quantisation to use. Select which of the following would provide a good enough sampling rate and split the signal into 9 levels.



- (a) Sampling frequency = 10Hz, quantisation = 4 bits
- (b) Sampling frequency = 8Hz, quantisation = 3 bits
- (c) Sampling frequency = 2Hz, quantisation = 4 bits
- (d) Sampling frequency = 2Hz, quantisation = 3 bits
- (e) None of the above

[2 marks]

2. A collaborator records the brain activity of an olympic athlete while running, and asks for your help to analyse it. You have already decided that the best sampling frequency is 2Hz and that you are going to use 3 bits to digitise the analog signal. What would be the correct binary representation of the signal plotted in the image below.



- (a) 000111010110010111
- (b) 001111010110010111
- (c) 001011010110010101
- (d) 001111010100010111
- (e) 000111011110010111

[3 marks]

3. Compute the 1-norm distance (L1) between these datapoints A and B. Where $A=(3,1,5,3,5,8,1,0)$ and $B=(2,1,3,5,5,8,1,1)$

- (a) $L1(A,B) = 6$
- (b) $L1(A,B) = 5$
- (c) $L1(A,B) = 7$
- (d) $L1(A,B) = 8$
- (e) None of the above

[2 marks]

4. You collected a four dimensional dataset of values $\mathbf{x} = (x_1, x_2, x_3, x_4)$ with means $(3, 2, 3.5, 2.6)$ and the following covariance matrix

$$\begin{bmatrix} 4 & 3 & -0.1 & -1 \\ 3 & 0.01 & -0.1 & 0 \\ -0.1 & -0.1 & 4 & 0.1 \\ -1 & 0 & 0.1 & 9 \end{bmatrix}$$

You are asked to select x_1 and one another variable to be processed by a machine learning algorithm. Which variable would you pick to best complement the information provided by x_1 ?

- (a) x_1
- (b) x_2
- (c) x_3
- (d) x_4
- (e) x_3 and x_4

[3 marks]

5. You are asked to find the most informative dimension of a 2D dataset with the following covariance matrix

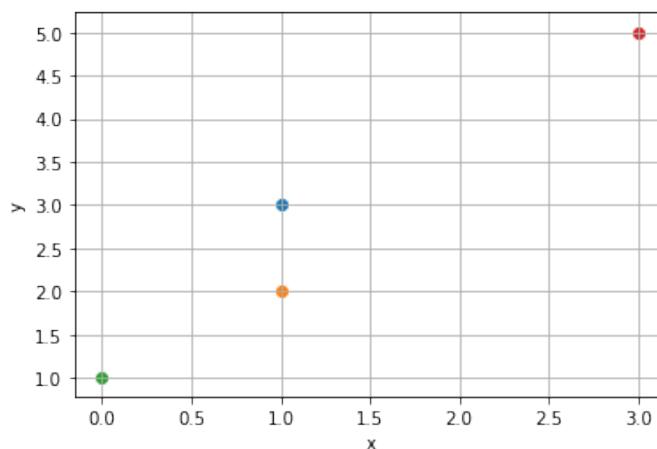
$$\begin{bmatrix} 3 & 0.2 \\ 0.2 & 5 \end{bmatrix}$$

What is the most informative eigenvalue and its respective eigenvector with unit vector length?

- (a) $\lambda \sim 3$ and $v = [\frac{1}{\sqrt{101}} \quad \frac{10}{\sqrt{101}}]$
- (b) $\lambda \sim 3$ and $v = [\frac{1}{10} \quad 1]$
- (c) $\lambda \sim 3$ and $v = [\frac{1}{2} \quad \sqrt{\frac{3}{4}}]$
- (d) $\lambda \sim 5$ and $v = [\frac{1}{\sqrt{101}} \quad \frac{10}{\sqrt{101}}]$
- (e) $\lambda \sim 5$ and $v = [\frac{1}{2} \quad \sqrt{\frac{3}{4}}]$

[4 marks]

6. You are given four 2D data points (x,y) represented in the figure below. Assuming a linear model of the form $y = a + bx$ use the general matrix form least squares method to tune the model's parameters. What is the result that you obtain:



- (a) $y = 1.2 + 1.4x$
- (b) $y = 1 + 2x$

- (c) $y = 1 + 2.3x$
 (d) $y = 1.1 + 1.3x$
 (e) $y = 1.1 + 2.1x$ [3 marks]

7. Using the multivariate normal distribution, calculate the probability of the datapoint $x = (1, -1.9)$ using mean $(1, -2)$ and covariance matrix $\begin{bmatrix} 3 & 0.2 \\ 0.2 & 5 \end{bmatrix}$.

- (a) 0.054
 (b) 0.032
 (c) 0.041
 (d) 0.011
 (e) 0.049 [3 marks]

8. For the data displayed in the table, compute the least-squares parameter fit for a model of the form, $\hat{y} = w_0 + w_1x$.

x	y
0.5	3.2
1.0	2.4
1.5	2.1
2.0	1.1

- (a) $\hat{y} = 2.89 - 0.76x$
 (b) $\hat{y} = 4.02 - 0.89x$
 (c) $\hat{y} = 3.57 - 1.29x$
 (d) $\hat{y} = 3.85 - 1.32x$
 (e) $\hat{y} = 3.82 - 1.36x$

[3 marks]

9. Given the dataset in the table, which model has the lowest squared-error?

x	y
-1.0	-0.4
0.0	2.4
1.0	4.3

- (a) $\hat{y} = 2$
 (b) $\hat{y} = 2x + 2$
 (c) $\hat{y} = 3x + 2$
 (d) $\hat{y} = -0.3x^2 + 2x + 2$
 (e) $\hat{y} = 0.2x^3 + 2x + 2$

[3 marks]

10. What issue must we consider in regression, but not in classification?

- (a) model-complexity
- (b) fitting the noise level, σ
- (c) overfitting
- (d) regularisation
- (e) cross-validation

[2 marks]

11. For the following dataset, fit Gaussian distributions to each class using maximum-likelihood, then compute the corresponding posterior over the class-label for $x = -0.3$ (specifically $P(y=1|x_0 = -0.3)$).

x	y
-1.0	0
-1.9	0
-1.5	0
3.1	1
1.3	1
0.7	1

- (a) 0.57
- (b) 0.78
- (c) 0.89
- (d) 0.97
- (e) 0.99

[3 marks]

12. In K-means, consider initializing the algorithm with two cluster centers at -1 and 1 , and data at,

x
-3.0
-2.3
-1.5
1.2
1.3
1.7

Compute updated cluster centers under a single K-means update step.

- (a) -2.02 and 1.46
- (b) -2.27 and 1.40
- (c) -2.73 and 1.53
- (d) -2.32 and 1.45

- (e) -2.42 and 1.32

[3 marks]

13. In a Gaussian mixture model, with two clusters,

$$\begin{aligned}P(z=0) &= P(z=1) = 0.5 \\P(x|z=0) &= \mathcal{N}(x; -1, 0.25) \\P(x|z=1) &= \mathcal{N}(x; 1, 2)\end{aligned}$$

Compute $P(z=0|x=-0.2)$.

- (a) 0.049
- (b) 0.023
- (c) 0.104
- (d) 0.054
- (e) 0.046

[3 marks]

14. The trigonometric Fourier series of an odd function of time does not contain the...

- (a) DC term
- (b) cos term
- (c) harmonic term
- (d) sine term
- (e) None of the above.

[2 marks]

15. The eigenvalues of a dataset are: [24, 19, 11, 4, 1, 0.88, 0.21]. Approximately, what variance in the dataset do the first 4 eigenvalues represent (when rounded)?

- (a) 97.0%
- (b) 96.5%
- (c) 96.0%
- (d) 95.5%
- (e) 95.0%

[3 marks]

16. Matrix K is a covariance matrix of some 3D data:

$$K = \begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix}$$

Which of the following sets is the correct eigenvalues and eigenvectors of K ?

- (a) Set A \longrightarrow $\lambda_A = 7$ and $\mathbf{e}_A = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$,
- (b) Set B \longrightarrow $\lambda_B = -2$ and $\mathbf{e}_B = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$,
- (c) Set C \longrightarrow $\lambda_C = -5$ and $\mathbf{e}_C = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$,
- (d) Set D \longrightarrow $\lambda_D = 1$ and $\mathbf{e}_D = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$,
- (e) Both Sets B and C

[4 marks]

17. Your friend was recently on an 11-hour, long-haul flight. Each hour during the flight, he or she made a note of the "outside air temperature" reading as displayed on the aircraft information screen. The readings are shown as signal R in Equation (1) below.

$$R = \begin{pmatrix} 0 & -25 & -44 & -70 & -70 & -50 & -50 & -40 & 0 \end{pmatrix}, \quad (1)$$

You decide to apply an averaging filter f to smooth this R signal, that then results in signal S (noting edge effects when performing convolution) in Equation (2) below, i.e. $S = R * f$:

$$S = \frac{1}{6} \begin{pmatrix} 0 & -50 & -138 & -278 & -368 & -380 & -340 & -280 & -180 & -80 & 0 \end{pmatrix}. \quad (2)$$

Given the $\frac{1}{6}$ normalisation factor, which of these filters below is the correct filter f ?

- (a) $h = \frac{1}{6}(-1 \ 4 \ -1)$
- (b) $h = \frac{1}{6}(1 \ 4 \ 1)$
- (c) $h = \frac{1}{6}(2 \ 2 \ 2)$
- (d) $h = \frac{1}{6}(1 \ 1 \ 1)$
- (e) $h = \frac{1}{6}(0 \ 5 \ 1)$

[3 marks]

18. Consider how the Fourier domain of a signal is affected under the following operations:
 (i) translation or delay shift of an audio signal, (ii) rotation of an object within an image on a uniform background. Which of these options is TRUE:

- (a) Under translation, the frequency magnitudes in the Fourier domain are shifted in a positive direction by the amount of translation. Under rotation, the Fourier domain magnitudes are rotated by an amount corresponding to the rotated object.
- (b) Under translation, the Fourier domain is not affected and the frequency magnitudes retain their position. Under rotation, the Fourier domain magnitudes also remain in the same position.
- (c) Under translation, the Fourier domain is not affected and the frequency magnitudes retain their position. Under rotation, the Fourier domain magnitudes are rotated by an amount corresponding to the rotated object.
- (d) Under translation, the frequency magnitudes in the Fourier domain are shifted in a negative direction to the translation. Under rotation, the Fourier domain magnitudes remain the in the same position.
- (e) None of the above.

[2 marks]

19. Consider the satellite image below (Moon Surface), of a small area of the surface of the Moon, which has suffered from some linear noise during transmission back to Earth. The image below-right (Cleaned-Image) is a cleaned-up version after the Fourier space of the original image was altered using a special mask. The second row of the figure shows 5 masks, labelled (A, B, C, D, E), one of which was used to produce the cleaned-up image.

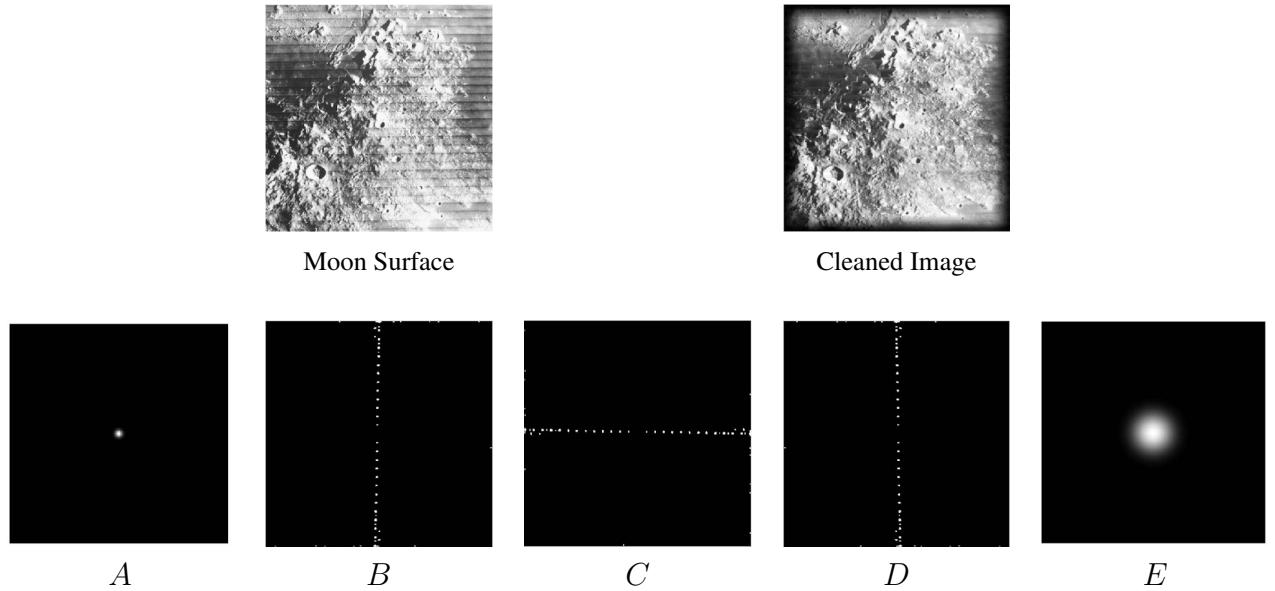


Figure 1: (top) Original Moon Surface image and its cleaned version, (bottom) Five possible masks applied to the Fourier space of the Moon Surface image.

Which of the above masks was used to clean the image?

- (a) Mask A
- (b) Mask B
- (c) Mask C
- (d) Mask D
- (e) Mask E

[6 marks]

END OF PAPER