Name:	Roll no.	A	U	2							Marks /20

Mark your answers on this sheet and solve questions sequentially in the answer book. Attach this sheet to the answer book and then submit it. Answer all questions. Marking Scheme: +2 for correct answer, -0.5 for incorrect answer. Tick the correct answer.

1. A dataset in $\mathbb{R}^4$ has the following covariance matrix C with eigenvalues:

$$\lambda_1 = 48, \lambda_2 = 24, \lambda_3 = 20, \lambda_4 = 8$$

A PCA is performed. You select the **first two principal components** and project all data onto the 2D subspace. A new test point $\mathbf{x} = [3,1,-2,4]^T$ is expressed in the **eigenvector basis** as:

$$\mathbf{z} = W^T\mathbf{x}, \text{ where } W = [\mathbf{w}_1\ \mathbf{w}_2\ \mathbf{w}_3\ \mathbf{w}_4]$$

with eigenvectors (columns of W):

$$\mathbf{w}_1 = \frac{1}{2}[1,1,1,1]^T, \mathbf{w}_2 = \frac{1}{2}[1,1,-1,-1]^T$$

$$\mathbf{w}_3 = \frac{1}{2}[1,-1,1,-1]^T, \mathbf{w}_4 = \frac{1}{2}[1,-1,-1,1]^T$$

Part (a): What is the **squared reconstruction error** $\|\mathbf{x} - \hat{\mathbf{x}}\|^2$ when only the first 2 principal components are used to reconstruct $\mathbf{x}$?

Part (b): What **percentage of total variance** is retained by these 2 components?

Pick the option where **both** answers are correct:

A) Reconstruction error = 10, Variance retained = 72%
B) Reconstruction error = 20, Variance retained = 72%
C) Reconstruction error = 10, Variance retained = 80%
D) Reconstruction error = 20, Variance retained = 80%

2 A binary classification problem has two classes with the following data:

- **Class 1** (3 samples): $\{[2,4]^T, [4,6]^T, [6,8]^T\}$
- **Class 2** (3 samples): $\{[1,1]^T, [3,3]^T, [5,5]^T\}$

Compute the class means μ_1 and μ_2, the **within-class scatter matrix** $S_W = S_1 + S_2$, and find the optimal LDA projection direction $w^* = S_W^{-1}(\mu_1 - \mu_2)$.

The **Fischer criterion** is: $J(w^*) = \frac{(w^{*T}(\mu_1 - \mu_2))^2}{w^{*T}S_W w^*}$

Given that S_W is singular, use regularization $S_W^{\text{reg}} = S_W + 0.5I$ where:

$$S_W = \begin{bmatrix} 16 & 16 \\ 16 & 16 \end{bmatrix}, S_W^{\text{reg}} = \begin{bmatrix} 16.5 & 16 \\ 16 & 16.5 \end{bmatrix}$$

What is the value of $J(w^*)$?

A) $J(w^*) = 4.25$
B) $J(w^*) = 16.50$
C) $J(w^*) = 24.05$
D) $J(w^*) = 32.80$

3 A logistic regression model with weight vector $\mathbf{w} = [0.5, -1.0]^T$ and **no bias** is trained on 3 samples:

Sample	x_i	y_i
1	[2,1]	1
2	[1,2]	0
3	[3,1]	1

Sigmoid: $\sigma(z) = \frac{1}{1+e^{-z}}$. Use: $e^{-1.5} \approx 0.223$, $e^{-0.5} \approx 0.607$

Compute predicted probabilities $\hat{y}_i = \sigma(\mathbf{w}^T x_i)$, then the gradient:

$$\nabla J = \frac{1}{3}\sum_{i=1}^{3} (\hat{y}_i - y_i)x_i$$

With learning rate $\eta = 1.0$, what is $\mathbf{w}^{\text{new}} = \mathbf{w} - \eta \nabla J$?

A) $\mathbf{w}^{\text{new}} \approx [1.15, -0.83]^T$
B) $\mathbf{w}^{\text{new}} \approx [0.78, -0.74]^T$
C) $\mathbf{w}^{\text{new}} \approx [0.50, -1.00]^T$
D) $\mathbf{w}^{\text{new}} \approx [0.65, -1.17]^T$

4 A 2D feature vector $\mathbf{x} = [4,0]^T$ is **rotated by 60° counter-clockwise**, then **projected onto the x-axis**.

Rotation matrix: $R(60°) = \begin{bmatrix} \cos 60° & -\sin 60° \\ \sin 60° & \cos 60° \end{bmatrix} = \begin{bmatrix} 0.5 & -0.866 \\ 0.866 & 0.5 \end{bmatrix}$

Projection matrix: $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Apply transformations sequentially: first rotate to get $\mathbf{x}'$, then project to get $\mathbf{x}''$.

What is the **squared norm** $\|\mathbf{x}''\|^2$?

A) 16
B) 4
C) 8
D) 12

5 Input $\mathbf{x} = [2,3]^T$ is mapped to degree-2 polynomial features with $\sqrt{2}$ scaling on cross-terms:

$$\phi(\mathbf{x}) = [1, \sqrt{2}x_1, \sqrt{2}x_2, x_1^2, x_2^2, \sqrt{2}x_1x_2]^T$$

A second point $\mathbf{x}' = [1,1]^T$ has $\phi(\mathbf{x}') = [1, \sqrt{2}, \sqrt{2}, 1, 1, \sqrt{2}]^T$.

Compute the **kernel value** (dot product in feature space):

$$k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \phi(\mathbf{x}')$$

Verify it equals $(1 + \mathbf{x}^T\mathbf{x}')^2 = (1 + 2 + 3)^2$.

What is $k(\mathbf{x}, \mathbf{x}')$?

A) $k = 25$
B) $k = 36$
C) $k = 49$ **D)** $k = 64$

6	A **centered** dataset has 4 data points in $\mathbb{R}^2$: $$X = \begin{bmatrix} 2 & 1 \\ -2 & -1 \\ 1 & 3 \\ -1 & -3 \end{bmatrix}$$ Compute covariance matrix $C = \frac{1}{4}X^T X$, find eigenvalues via $\det(C - \lambda I) = 0$. Given $C = \begin{bmatrix} 2.5 & 2.5 \\ 2.5 & 5.0 \end{bmatrix}$, the characteristic equation is: $$\lambda^2 - 7.5\lambda + 6.25 = 0$$ What is the **proportion of variance explained** by the first principal component (largest eigenvalue)? **A)** 70.0% **B)** 80.0% **C)** 90.0% **D)** 95.0%
7	Linear transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$. Observed output: $\mathbf{b} = [7,13]^T$. Recover input $\mathbf{x} = A^{-1}\mathbf{b}$. Additionally, compute $\det(B)$ where $B = A^2 - 3A + I$. Matrix A has eigenvalues satisfying $\lambda^2 - 5\lambda + 1 = 0$, so $\lambda_1\lambda_2 = 1$. If λ is eigenvalue of A, then $\lambda^2 - 3\lambda + 1$ is eigenvalue of B. What is $\mathbf{x}$ and $\det(B)$? **A)** $\mathbf{x} = [1,4]^T$, $\det(B) = 4$ **B)** $\mathbf{x} = [3,2]^T$, $\det(B) = 0$ **C)** $\mathbf{x} = [1,4]^T$, $\det(B) = 1$ **D)** $\mathbf{x} = [3,2]^T$, $\det(B) = 4$
8	A binary classifier evaluated on $n = 10$ independent samples gets $k = 7$ correct predictions. Log-likelihood: $\ell(p) = \log \binom{n}{k} + k\log p + (n-k)\log(1-p)$ Taking $\frac{d\ell}{dp} = 0$ gives MLE: $\hat{p} = \frac{k}{n}$ Fisher Information per sample: $\mathcal{I}_{\text{per sample}}(p) = \frac{1}{p(1-p)}$ Variance of MLE: $\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$ What is $\hat{p}_{\text{MLE}}$ and $\text{Var}(\hat{p})$ at $\hat{p} = 0.7$? **A)** $\hat{p} = 0.7$, $\text{Var}(\hat{p}) = 0.021$ **B)** $\hat{p} = 0.7$, $\text{Var}(\hat{p}) = 0.035$ **C)** $\hat{p} = 0.5$, $\text{Var}(\hat{p}) = 0.025$ **D)** $\hat{p} = 0.7$, $\text{Var}(\hat{p}) = 0.049$

9	LDA optimization: $\max_{\mathbf{w}} J(\mathbf{w}) = \frac{\mathbf{w}^T S_B \mathbf{w}}{\mathbf{w}^T S_W \mathbf{w}}$ Scatter matrices: $S_W = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$, $S_B = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$ Lagrangian with constraint $\mathbf{w}^T S_W \mathbf{w} = 1$ gives: $S_B \mathbf{w} = \lambda S_W \mathbf{w}$ Compute $S_W^{-1} S_B = \begin{bmatrix} 2.25 & 0 \\ 0 & 4 \end{bmatrix}$ (diagonal, eigenvalues are 2.25 and 4.0) Maximum Fischer criterion $\lambda^* =$ max eigenvalue. For $\lambda^* = 4.0$, eigenvector is $\mathbf{w}^* = [0,1]^T$ (ratio 0: 1). What is λ^* and the eigenvector ratio? **A)** $\lambda^* = 4.0$, ratio 0: 1 **B)** $\lambda^* = 2.25$, ratio 1: 0 **C)** $\lambda^* = 4.0$, ratio 1: 2 **D)** $\lambda^* = 2.25$, ratio 2: 1
10	Four vectors in $\mathbb{R}^3$: $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$ Observe: $\mathbf{v}_2 = 2\mathbf{v}_1$ (linearly dependent), $\mathbf{v}_4 = \mathbf{v}_1 + 2\mathbf{v}_3$ (linearly dependent). Maximal independent subset: $\{\mathbf{v}_1, \mathbf{v}_3\}$ with dim (span) = 2. Target vector $\mathbf{b} = [5,4,9]^T$ expressed as $\mathbf{b} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_3$. System: $\alpha + \beta = 5$, $2\alpha = 4$, $3\alpha + \beta = 9$ What is dim (span) and (α, β)? **A)** dim = 3, $\alpha = 2, \beta = 3$ **B)** dim = 2, $\alpha = 2, \beta = 3$ **C)** dim = 2, $\alpha = 3, \beta = 2$ **D)** dim = 3, $\alpha = 3, \beta = 2$