DO NOT WRITE YOUR NAME OR KAUST ID NUMBER ON THIS PAGE OR ANY OTHER PAGE

PUT YOUR EXAM ID NUMBER ON THIS PAGE AND EVERY OTHER PAGE YOU SUBMIT

WRITE YOUR SOLUTIONS ON ONLY ONE SIDE OF EMPTY SOLUTION SHEETS PROVIDED

PUT PAGE NUMBERS ON ALL SOLUTION PAGES

YOUR SOLUTIONS SHOULD BE ORGANIZED WELL AND WRITTEN CLEARLY – NEAT AND EASY-TO-READ SOLUTIONS WILL HELP YOU IN GRADING

DO NOT WRITE ANY PART OF YOUR SOLUTIONS ON PROBLEM SHEETS – SOLUTIONS ON PROBLEM SHEETS WILL NOT BE GRADED

YOU ARE ALLOWED TO SUBMIT SOLUTIONS TO ONLY FIVE PROBLEMS

EXAM ID NUMBER:________

COURSE NUMBER: AMCS 211

PROBLEM: 1

Principal Component Analysis [20 pts]

In this problem, we will derive a closed form solution to a traditional problem in pattern recognition, called principal component analysis (PCA).

1. To simplify the discussion of PCA, consider the constrained optimization problem in Eq (1), where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric.

$$\begin{aligned} \mathbf{x}^* = \ & \arg\max_{\mathbf{x}} \ f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \\ & \text{subject to: } \|\mathbf{x}\|_2^2 = 1 \end{aligned} \tag{1}$$

   (i) **[3 pts]** Write down the KKT conditions for Eq (1).

   (ii) **[3 pts]** Show that $\mathbf{x}^*$ must be an eigenvector of matrix $\mathbf{A}$.

   (iii) **[4 pts]** Show that $f(\mathbf{x}^*) = \sigma_1(\mathbf{A})$, where $\sigma_1(\mathbf{A})$ is the maximum eigenvalue of $\mathbf{A}$ and $\mathbf{x}^*$ is its corresponding eigenvector.

2. Consider a set of m points $\{\mathbf{y}_i\}_{i=1}^m$, where each $\mathbf{y}_i \in \mathbb{R}^n$. We seek the unit norm direction $\mathbf{w} \in \mathbb{R}^n$, along which the empirical variance of the projection of all m points is *maximum*. Here, we define the projection of $\mathbf{y}_i$ along $\mathbf{w}$ as: $p_i = \mathbf{w}^T \mathbf{y}_i$. We can form the empirical variance of this projection as in Eq (2).

$$h(\mathbf{w}) = \frac{1}{m} \sum_{i=1}^{m} (p_i - \bar{p})^2, \quad \text{where } \bar{p} = \frac{1}{m} \sum_{i=1}^{m} p_i \tag{2}$$

   (i) **[6 pts]** Let $\bar{\mathbf{y}} = \frac{1}{m} \sum_{i=1}^m \mathbf{y}_i$ and $\mathbf{z}_i = \mathbf{y}_i - \bar{\mathbf{y}}$. Show that: $h(\mathbf{w}) = \mathbf{w}^T \mathbf{C} \mathbf{w}$, where $\mathbf{C} = \frac{1}{m} \mathbf{Z}\mathbf{Z}^T$ is the outer product matrix of the m points and $\mathbf{z}_i$ is the i^{th} column of matrix $\mathbf{Z} \in \mathbb{R}^{n \times n}$.

   (ii) **[4 pts]** PCA aims to find a unit ℓ_2 norm vector $\mathbf{w}$ that maximizes the empirical projection variance $h(\mathbf{w})$. Formulate PCA as an optimization problem and use the previous results to find a closed form global solution.

DO NOT WRITE YOUR NAME OR KAUST ID NUMBER ON THIS PAGE OR ANY OTHER PAGE

PUT YOUR EXAM ID NUMBER ON THIS PAGE AND EVERY OTHER PAGE YOU SUBMIT

WRITE YOUR SOLUTIONS ON ONLY ONE SIDE OF EMPTY SOLUTION SHEETS PROVIDED

PUT PAGE NUMBERS ON ALL SOLUTION PAGES

YOUR SOLUTIONS SHOULD BE ORGANIZED WELL AND WRITTEN CLEARLY – NEAT AND EASY-TO-READ SOLUTIONS WILL HELP YOU IN GRADING

DO NOT WRITE ANY PART OF YOUR SOLUTIONS ON PROBLEM SHEETS – SOLUTIONS ON PROBLEM SHEETS WILL NOT BE GRADED

YOU ARE ALLOWED TO SUBMIT SOLUTIONS TO ONLY FIVE PROBLEMS

EXAM ID NUMBER:_______

COURSE NUMBER: AMCS 211

PROBLEM: 2

The Support Vector Machine as a Quadratic Program (QP) [20 pts]

Support Vector Machines (SVM) are widely used in pattern recognition and classification. The parameter $\mathbf{x}$ of an SVM is computed by solving the optimization problem in Eq (3).

$$\min_{\mathbf{x}} \ \frac{1}{2}\|\mathbf{x}\|_2^2 + \frac{1}{n}\sum_{i=1}^{n} \max(0, 1 - \mathbf{x}^T\mathbf{a}_i) \tag{3}$$

1. **[5 pts]** Show that Eq (3) can be formulated as a convex QP
2. **[5 pts]** Write down the corresponding Lagrangian and the first-order KKT conditions
3. **[10 pts]** Derive the dual optimization problem of Eq (3) and show that it can be formulated as a convex QP with box constraints as in Eq (4). Determine $\mathbf{G}$, $\mathbf{c}$, $\mathbf{l}$, and $\mathbf{u}$.

$$\begin{aligned} \min_{\mathbf{z}} \ & \frac{1}{2}\mathbf{z}^T\mathbf{G}\mathbf{z} + \mathbf{c}^T\mathbf{z} \\ \text{subject to: } & \mathbf{l} \le \mathbf{z} \le \mathbf{u} \end{aligned} \tag{4}$$