Exercise 9:

1 (a)

The Primal form of the One-class SVM has the form

$$\min_{R, c, (\xi_i)_{i=1}^{n}} R^2 + \frac{1}{nv} \sum_{i=1}^{n} \xi_i$$

such that

$$\forall_{i=1}^{n} : \|\phi(x_i) - c\|^2 \leq R^2 + \xi_i \text{ and } \xi_i \geq 0.$$

Using Lagrange's

$$\min_{R, c, \xi} \max_{\alpha, \beta \geq 0} \underbrace{\left\{ R^2 + \frac{1}{nv} \sum_{i=1}^{N} \xi_i + \sum_{i=1}^{N} \alpha_i \|\phi(x_i) - c\|^2 - R^2 - \xi_i - \sum_{i=1}^{N} \beta_i \xi_i \right\}}_{L(R, c, \xi, \alpha, \beta)}$$

Dual optimization problem is given as

$$\max_{\alpha, \beta \geq 0} g(\alpha, \beta)$$

$$g(\alpha, \beta) = \min_{R, C, \xi} L(R, C, \xi, \alpha, \beta)$$

Taking Partial derivative of L w.r.t c, R, ξ_i

$$\frac{\partial L}{\partial c} = \sum_{i=1}^{N} \alpha_i \cdot 2(\phi(x_i) - c) = 0 \rightarrow c \sum_{i=1}^{N} \alpha_i = \sum_{i=1}^{N} \alpha_i \phi(x_i) \quad \text{....(1)}$$

$$\frac{\partial L}{\partial R} = 2R - 2\sum_{i=1}^{N} \alpha_i R = 0 \Rightarrow \left(1 - \sum_{i=1}^{N} \alpha_i\right) R = 0. \Rightarrow \sum_{i=1}^{N} \alpha_i = 1 \quad \text{...(2)}$$

$$\frac{\partial L}{\partial \xi_i} = \frac{1}{nv} - \alpha_i - \beta_i = 0 \Rightarrow \alpha_i + \beta_i = 1/nv \quad \text{...(3)}$$

substituting (2) in (1) give

$$c = \sum_{i=1}^{N} \alpha_i \phi(x_i)$$

$g(\alpha, \beta)$ can be written as

$$g(\alpha, \beta) = R^2 + \sum_{i=1}^{N} (\alpha_i + \beta_i) \xi_i + \sum_{i=1}^{N} \alpha_i \left(\|\phi(x_i) - \sum_{i=1}^{N} \alpha_i \phi(x_i)\|^2 - R^2 - \xi_i \right) - \sum_{i=1}^{N} \beta_i \xi_i$$

$$= R^2 - R^2 \sum_{i=1}^{N} \alpha_i + \sum_{i=1}^{N} \alpha_i \left((\phi(x_i) - \sum_{j=1}^{N} \alpha_j \phi(x_j))^T (\phi(x_i) - \sum_{j=1}^{N} \alpha_j \phi(x_j)) \right)$$

using $\sum_{i=1}^{N} \alpha_i = 1$

$$g(\alpha, \beta) = \sum_{i=1}^{N} \alpha_i \phi(x_i)^T \phi(x_i) - \sum_{i,j=1}^{N} \alpha_i \alpha_j \phi(x_i)^T \phi(x_j)$$

Lets say we have kernel function k

$$k(x,y) = \Phi(x)^T \Phi(y)$$

$$\max_{\alpha} \left\{ \sum_{i=1}^{N} \alpha_i k(x_i, x_i) - \sum_{i,j=1}^{N} \alpha_i \alpha_j k(x_i, x_j) \right\}$$

with constraint $\sum_{i=1}^{N} \alpha_i = 1$ and $c = \sum_{i=1}^{N} \alpha_i \Phi(x_i)$

Ex 2: matrix form: $\min\limits_{\alpha} \alpha^T P \alpha + q^T \alpha$ s.t. $G\alpha \le h$ & $A\alpha = b$

dual program: $\min\limits_{\alpha} \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \underbrace{k(x_i, x_j)}_{P} \alpha_j + \sum_{i=1}^{n} \underbrace{k(x_i, x_i)}_{q^T} \alpha_i$ s.t. $\overset{G}{1} \cdot \alpha_i \le \overset{h}{\frac{1}{n\nu}}$ & $\overset{A}{1} \cdot \sum \alpha_i = \overset{b}{1}$

& $c = \sum_{i=1}^{n} \alpha_i \phi(x_i)$

$P \in \mathbb{R}^{n \times n}$

$G \in \mathbb{R}^{2n \times n}$ $\quad \to \underbrace{\begin{pmatrix} 1 & & \\ & \ddots & 1 \\ -1 & \ddots & \\ & & -1 \end{pmatrix}}_{G} \cdot \underbrace{\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}}_{\alpha} \le \underbrace{\begin{pmatrix} \frac{1}{n\nu} \\ \vdots \\ 0 \end{pmatrix}}_{h} \quad \to 0 \le \alpha_i \le \frac{1}{n\nu}$

$A \in \mathbb{R}^{2 \times n}$ $\quad \to \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ \phi(x_1) & \cdots & \phi(x_n) \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}}_{\alpha} = \underbrace{\begin{pmatrix} 1 \\ c \end{pmatrix}}_{b} \quad \to \sum_{i=1}^{n} \alpha_i = 1$ & $c = \sum_{i=1}^{n} \alpha_i \phi(x_i)$

$q \in \mathbb{R}^{n \times 1}$

$h \in \mathbb{R}^{2n \times 1}$

$b \in \mathbb{R}^{2 \times 1}$