1 a)

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

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

$$L(R, c, \xi_i, \lambda, \eta) = R^2 + \frac{1}{n\nu}\sum \xi_i - \sum \lambda_i \left(R^2 + \xi_i - \|\phi(x_i) - c\|^2\right) - \sum \eta_i \xi_i$$

$$c_1(R, c, \xi_i, x_i) = R^2 + \xi_i - \|\phi(x_i) - c\|^2 (\geq 0)$$

$$c_2(\xi_i) = \xi_i \ (\geq 0)$$

$$\frac{\partial L}{\partial R} = 2R - 2R\sum\lambda_i = 0 \qquad \sum \lambda_i = 1$$

$$\frac{\partial L}{\partial c} = -2\sum\lambda_i(\phi(x_i) - c) = 0 \qquad c = \frac{\sum \lambda_i \phi(x_i)}{\sum \lambda_i} = \sum \lambda_i \phi(x_i)$$

$$\frac{\partial L}{\partial \xi_i} = \frac{1}{n\nu} - \lambda_i - \eta_i = 0 \qquad \eta_i = \frac{1}{n\nu} - \lambda_i$$

Since $\eta_i \geq 0$ we can take this term away from expression and get inequality constraints

$$0 \leq \lambda_i \leq \frac{1}{n\nu}$$

For the dual problem we substitute expressions for some variables derived above and apply derived constraints:

$$L(\lambda_i) = R^2 + \frac{1}{n\nu}\sum_i^n \xi_i - \sum_i^n \lambda_i\left(R^2 + \xi_i - \left(\phi(x_i) - \sum_j^n \lambda_j \phi(x_j)\right)^2\right) - \sum_i^n \left(\frac{1}{n\nu} - \lambda_i\right)\xi_i$$

$$(1) \quad = R^2 + \frac{1}{n\nu}\sum_i^n \xi_i - R^2\sum_i^n \lambda_i + \sum_i^n \lambda_i \xi_i + \sum_i^n \lambda_i \Big(\phi(x_i)\phi(x_i) - \phi(x_i)\sum_j^n \lambda_j \phi(x_j) - $$
$$- \phi(x_i)\sum_j^n \lambda_j \phi(x_j) + \sum_j^n \lambda_j \phi(x_j) \sum_k^n \lambda_k \phi(x_k)\Big) - \frac{1}{n\nu}\sum \xi_i + \sum \lambda_i \xi_i =$$

$$(2) = \sum_i^n \lambda_i \phi(x_i)\phi(x_i) - 2\sum_i^n \lambda_i \phi(x_i) \sum_j^n \lambda_j \phi(x_j) + \sum_j^n \lambda_j \phi(x_j) \sum_k^n \lambda_k \phi(x_k) \sum \lambda_i =$$

$$(3) = \sum_i^n \lambda_i \phi(x_i)\phi(x_i) - \sum_i^n \sum_j^n \lambda_i \lambda_j \phi(x_i)\phi(x_j)$$

$$\text{s.t. } \sum \lambda_i = 1 \quad \text{and} \quad 0 \leq \lambda_i \leq \frac{1}{n\nu}$$

where for (1) we substituted derived expressions for c and η_i and for (2) we took advantage of constraint $\sum \lambda_i = 1$ and eliminated some terms.

Finally, we got (3) be using the fact that two last terms in (2) are identical since they both iterate over the same set of values. Also, since the last term didn't depend on index i, we could factor out $\sum \lambda_i$ and apply constraint $\sum \lambda_i = 1$.

b) To kernelize dual we can replace all dot products with kernels: $\phi(x)\phi(x') \mapsto k(x,x')$

The new objective becomes

$$\max_{\alpha} \sum_i^n \lambda_i \phi(x_i)\phi(x_i) - \sum_i^n \sum_j^n \lambda_i \lambda_j \phi(x_i)\phi(x_j)$$

$$= \sum_i^n \lambda_i k(x_i, x_i) - \sum_i^n \sum_j^n \lambda_i \lambda_j k(x_i, x_j)$$

$$\text{s.t. } \sum_i^n \lambda_i = 1$$

$$\forall_{i=1}^n : 0 \leq \lambda_i \leq \frac{1}{nv}$$

$$c = \sum_i^n \lambda_i \phi(x_i)$$

.2 Above we had maximization problem, so to make it a minimization problem to fit the form of a quadratic program, we put a negative sign before it.

New objective:

$$\min \sum_i^n \sum_j^n \lambda_i \lambda_j k(x_i x_j) - \sum_i^n \lambda_i k(x_i x_i)$$

We define $\lambda^T = [\lambda_1, \lambda_2, \dots, \lambda_n] \in \mathbb{R}^n$,

$$P = \begin{pmatrix} k(x_1, x_1) & k(x_1 x_2) & \dots & k(x_1 x_n) \\ & \vdots & & \\ k(x_n x_1) & k(x_n x_2) & \dots & k(x_n x_n) \end{pmatrix} \in \mathbb{R}^{n \times n},$$

and $q = [-k(x_1, x_1) \ -k(x_2 x_2) \ \dots \ -k(x_n x_n)] \in \mathbb{R}^n$.

Equality constraints we ~~transfor~~ write as

$$\sum_i^n \alpha_i = 1 \longmapsto \mathbb{1}^T \alpha = 1 \qquad A\alpha = b$$

$$A = \mathbb{1}^T \in \mathbb{R}^n$$

$$\alpha \in \mathbb{R}^n$$

$$b \in \mathbb{R}$$

For inequality constraints to fit the form $G\alpha \leq h$ we have to first split them:

$$0 \leq \alpha_i \leq \frac{1}{n\nu} \longmapsto \begin{array}{l} -\alpha_i \leq 0 \\ \alpha_i \leq \frac{1}{n\nu} \end{array} \quad \forall_{i=1}^n$$

To write it in matrix-vector form, we define G as

$$G = \begin{pmatrix} A \\ B \end{pmatrix}, \text{ where } A = \begin{pmatrix} -1 & 0 & 0 & \cdots \\ 0 & -1 & 0 & \\ \vdots & & \ddots & \\ & & & -1 \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad B = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ \vdots & & & \ddots \ 1 \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

The dimensionality of G is thus $2n \times n$, i.e. $G \in \mathbb{R}^{2n \times n}$.

Vector h should have dimensionality $2n$, where first n elements are zeros, and last n elements are $\frac{1}{n\nu}$.

Thus, $P \in \mathbb{R}^{n \times n}$, $\alpha \in \mathbb{R}^n$, $q \in \mathbb{R}^n$,

$A \in \mathbb{R}^n$, $b \in \mathbb{R}$

$G \in \mathbb{R}^{2n \times n}$, $h \in \mathbb{R}^{2n}$.