i) Discussed in class.

ii) Primal:

$$\min \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_i \xi_i^2$$

$$\text{s.t. } z_i(y)(\langle w, \phi(x_i, y)\rangle + \theta_i) \geq 1 - \xi_i$$
$$\xi_i \geq 0 \quad \forall i = 1, \dots, n, \ \forall y \in \mathcal{Y}$$

Build Lagrangian:

$$L = \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{n} \xi_i^2$$
$$- \sum_{i=1}^{n}\sum_{y} \alpha_i(y)\Big(z_i(y)\big(w^T\phi(x_i, y) + \theta_i\big) - 1 + \xi_i\Big)$$

with $\alpha_i(y) \geq 0$ $\begin{cases} \forall i = 1, \dots, n \\ \forall y \in \mathcal{Y} \end{cases}$ || since

$\leadsto$ max wrt. α and min wrt. w, ξ

$$\frac{\partial L}{\partial \xi_i} = C\xi_i - \sum_y \alpha_i(y) \overset{!}{=} 0$$

$$\Rightarrow \quad \xi_i = \frac{1}{C}\sum_y \alpha_i(y) \geq 0$$

$\leadsto$ answers question, why we do not need to enforce positivity of ξ_i ...

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{n}\sum_{y} \alpha_i(y)\, z_i(y)\, \phi(x_i, y) \overset{!}{=} 0$$

$$\Rightarrow w = \sum_{i=1}^{n} \sum_{y} \alpha_i(y) z_i(y) \phi(x_i, y)$$

$$\frac{\partial L}{\partial \theta_i} = -\sum_{y} \alpha_i(y) z_i(y) \stackrel{!}{=} 0$$

$$\Rightarrow \quad \sum_{y} \alpha_i(y) z_i(y) = 0$$

Further (KKT conditions):

$$\alpha_i(y) z_i(y) \left(w^T \phi(x_i, y) + \theta_i \right) - 1 + \xi_i = 0 \quad \forall i = 1, \dots, m$$

$$\alpha_i(y) \geq 0 \qquad \forall i = 1, \dots, n, \ \forall y \in \mathcal{Y}$$

$\rightsquigarrow$ plugging back into primal Lagrangian:

$$L = \frac{1}{2} \left\| \sum_{i=1}^{n} \sum_{y} \alpha_i(y) z_i(y) \phi(x_i, y) \right\|^2 + \frac{C}{2} \sum_{i=1}^{n} \left(\frac{1}{C} \sum_{y} \alpha_i(y) \right)^2$$

$$- \sum_{i=1}^{n} \sum_{y} \alpha_i(y) \Big(z_i(y) \Big(\underbrace{\Big[\sum_{j=1}^{n} \sum_{\bar{y}} \alpha_j(\bar{y}) z_j(\bar{y}) \phi(x_j, \bar{y}) \Big]}_{w}^T \phi(x_i, y) + \theta_i \Big) - 1 + \underbrace{\Big[\frac{1}{C} \sum_{\bar{y}} \alpha_i(\bar{y}) \Big]}_{\xi_i} \Big)$$

$$= \frac{1}{2} \Big(\sum_{i=1}^{n} \sum_{y} \sum_{j=1}^{n} \sum_{\bar{y}} \alpha_i(y) \alpha_j(\bar{y}) z_i(y) z_j(\bar{y}) k_{ij}(y, \bar{y})$$

$$+ \frac{1}{2C} \sum_{i=1}^{n} \sum_{y} \sum_{\bar{y}} \alpha_i(y) \alpha_i(\bar{y})$$

$$- \sum_{i=1}^{n} \sum_{y} \sum_{j=1}^{n} \sum_{\bar{y}} \alpha_i(y) \alpha_j(y) z_i(y) z_i(\bar{y}) k_{ij}(y, \bar{y}) \underbrace{- \sum_{i=1}^{n} \theta_i \underbrace{\sum_{y} \alpha_i(y) z_i(y)}_{=0}}_{=0}$$

$$+ \sum_{i} \sum_{y} \alpha_i(y) - \frac{1}{C} \sum_{i=1}^{n} \sum_{y} \sum_{\bar{y}} \alpha_i(y) \alpha_i(\bar{y})$$

$$= -\frac{1}{2} \Big(\sum_{i=1}^{n} \sum_{y} \sum_{j=1}^{n} \sum_{\bar{y}} \alpha_i(y) \alpha_j(\bar{y}) z_i(y) z_j(\bar{y}) k_{ij}(y, \bar{y})$$

$$\boxed{- \frac{1}{2C} \sum_{i=1}^{n} \sum_{y} \sum_{\bar{y}} \alpha_i(y) \alpha_i(\bar{y})} + \sum_{i} \sum_{y} \alpha_i(y)$$

$\boxed{}$ with additional term not in Altun et al. Eq. 16

with $k_{ij}(y, \bar{y}) = \langle \phi(x_i, y), \phi(x_j, \bar{y}) \rangle$

$\rightsquigarrow$ max wrt α !

iii) If we define

$$k^c_{ii}(y, \bar{y}) = \langle \phi(x_i, y), \phi(x_i, \bar{y}) \rangle + \frac{1}{C} z_i(y) z_i(y')$$

$$\forall i \neq j: \; k^c_{ij}(y, \bar{y}) = k_{ij}(y, \bar{y})$$

we can write the dual as

$$\underbrace{-\frac{1}{2}\left(\sum_{i=1}^{n} \sum_{y} \sum_{j=1}^{n} \sum_{\bar{y}} \alpha_i(y) \alpha_j(\bar{y}) z_i(y) z_j(\bar{y}) K^C_{ij}(y, \bar{y})\right)}_{A} + \sum_{i}^{n} \sum_{y} \alpha_i(y)$$

because all terms in A with $i=j$ sum up to

$$-\frac{1}{2} \sum_{i=1}^{n} \sum_{y} \sum_{\bar{y}} \alpha_i(y) \alpha_i(\bar{y}) z_i(y) z_i(\bar{y}) \left(k_{ii}(y, \bar{y}) + \frac{1}{C} z_i(y) z_i(y') \right)$$

$$= \underbrace{-\frac{1}{2} \sum_{i=1}^{n} \sum_{y} \sum_{\bar{y}} \alpha_i(y) \alpha_i(\bar{y}) z_i(y) z_i(\bar{y}) K_{ii}(y, \bar{y})}_{=B}$$

$$- \frac{1}{2} \sum_{i=1}^{n} \sum_{y} \sum_{\bar{y}} \alpha_i(y) \alpha_i(\bar{y}) \underbrace{\left(z_i(y) z_i(\bar{y})\right)^2}_{=1} \cdot \frac{1}{C}$$

$$= B \; \boxed{-\frac{1}{2C} \sum_{i=1}^{n} \sum_{y} \sum_{\bar{y}} \alpha_i(y) \alpha_i(\bar{y})} \quad \blacksquare$$

recovering the additional term from above