Exercise H10.1: The dual problem of the ν-SVR

(a)
$$\min_{\underline{w}, b, \varepsilon, \xi_\alpha, \xi_\alpha^*} \quad \frac{1}{2} \|\underline{w}\|^2 + C \left[\nu \varepsilon + \frac{1}{p} \sum_{\alpha=1}^{p} (\xi_\alpha + \xi_\alpha^*) \right]$$

s.t.
$$\underline{w}^T \underline{x}^{(\alpha)} + b - y_T^{(\alpha)} \leq \varepsilon + \xi_\alpha \qquad \forall \alpha \qquad \rightarrow \lambda_\alpha \geq 0$$
$$y_T^{(\alpha)} - \underline{w}^T \underline{x}^{(\alpha)} - b \leq \varepsilon + \xi_\alpha^* \qquad \rightarrow \lambda_\alpha^* \geq 0$$
$$\xi_\alpha \geq 0 \qquad \rightarrow \mu_\alpha \geq 0$$
$$\xi_\alpha^* \geq 0 \qquad \rightarrow \mu_\alpha^* \geq 0$$
$$\varepsilon \geq 0 \qquad \rightarrow \delta \geq 0$$

$$\mathcal{L}(\underline{w}, b, \varepsilon, \xi_\alpha, \xi_\alpha^*, \lambda_\alpha, \lambda_\alpha^*, \mu_\alpha, \mu_\alpha^*, \delta)$$
$$= \frac{1}{2} \|\underline{w}\|^2 + C \left[\nu \varepsilon + \frac{1}{p} \sum_{\alpha=1}^{p} (\xi_\alpha + \xi_\alpha^*) \right]$$
$$+ \sum_{\alpha=1}^{p} \lambda_\alpha \left(\underline{w}^T \underline{x}^{(\alpha)} + b - y_T^{(\alpha)} - \varepsilon - \xi_\alpha \right)$$
$$+ \sum_{\alpha=1}^{p} \lambda_\alpha^* \left(y_T^{(\alpha)} - \underline{w}^T \underline{x}^{(\alpha)} - b - \varepsilon - \xi_\alpha^* \right)$$
$$- \sum_{\alpha=1}^{p} \mu_\alpha \xi_\alpha$$
$$- \sum_{\alpha=1}^{p} \mu_\alpha^* \xi_\alpha^*$$
$$- \delta \varepsilon$$

$$\frac{\partial \mathcal{L}}{\partial \underline{w}} = \underline{w} + \sum_{\alpha=1}^{p} (\lambda_\alpha - \lambda_\alpha^*) \underline{x}^{(\alpha)} \stackrel{!}{=} 0 \qquad \boxed{\underline{w} = \sum_{\alpha=1}^{p} (\lambda_\alpha^* - \lambda_\alpha) \underline{x}^{(\alpha)}}$$

$$\frac{\partial \mathcal{L}}{\partial b} = \sum_{\alpha=1}^{p} (\lambda_\alpha - \lambda_\alpha^*) \stackrel{!}{=} 0 \qquad \boxed{\sum_{\alpha=1}^{p} (\lambda_\alpha - \lambda_\alpha^*) = 0}$$

$$\frac{\partial \mathcal{L}}{\partial \varepsilon} = C\nu - \sum_{\alpha=1}^{p} (\lambda_\alpha + \lambda_\alpha^*) - \delta \stackrel{!}{=} 0 \qquad \boxed{C\nu = \delta + \sum_{\alpha=1}^{p} (\lambda_\alpha + \lambda_\alpha^*)}$$

$$\frac{\partial \mathcal{L}}{\partial \xi_\alpha} = \frac{C}{p} - \lambda_\alpha - \mu_\alpha \stackrel{!}{=} 0 \qquad \boxed{C = p(\lambda_\alpha + \mu_\alpha)}$$

$$\frac{\partial \mathcal{L}}{\partial \xi_\alpha^*} = \frac{C}{p} - \lambda_\alpha^* - \mu_\alpha^* \qquad \boxed{C = p(\lambda_\alpha^* + \mu_\alpha^*)}$$

(b) $$\mathcal{L} = \frac{1}{2}\sum_{\alpha,\beta=1}^{p} (\lambda_\alpha^* - \lambda_\alpha)(\lambda_\beta^* - \lambda_\beta)(\underline{x}^{(\alpha)})^T \underline{x}^{(\beta)} + C\nu\varepsilon + \frac{C}{p}\sum_{\alpha=1}^{p}(\varphi_\alpha + \varphi_\alpha^*)$$
$$+ \sum_{\alpha,\beta=1}^{p} \lambda_\alpha (\lambda_\beta^* - \lambda_\beta)(\underline{x}^{(\alpha)})^T \underline{x}^{(\beta)} + b\sum_{\alpha=1}^{p}\lambda_\alpha - \sum_{\alpha=1}^{p}\lambda_\alpha y_T^{(\alpha)}$$
$$- \varepsilon\sum_{\alpha=1}^{p}\lambda_\alpha - \sum_{\alpha=1}^{p}\lambda_\alpha\varphi_\alpha + \sum_{\alpha=1}^{p}\lambda_\alpha^* y_T^{(\alpha)} - \sum_{\alpha,\beta=1}^{p}\lambda_\alpha^*(\lambda_\beta^* - \lambda_\beta)(\underline{x}^{(\alpha)})^T \underline{x}^{(\beta)}$$
$$- b\sum_{\alpha=1}^{p}\lambda_\alpha^* - \varepsilon\sum_{\alpha=1}^{p}\lambda_\alpha^* - \sum_{\alpha=1}^{p}\lambda_\alpha^*\varphi_\alpha^* - \frac{C}{p}\sum_{\alpha=1}^{p}\varphi_\alpha + \sum_{\alpha=1}^{p}\lambda_\alpha\varphi_\alpha$$
$$- \frac{C}{p}\sum_{\alpha=1}^{p}\varphi_\alpha^* + \sum_{\alpha=1}^{p}\lambda_\alpha^*\varphi_\alpha^* - C\nu\varepsilon + \varepsilon\sum_{\alpha=1}^{p}(\lambda_\alpha + \lambda_\alpha^*)$$

$$\mathcal{L} = \frac{1}{2}\sum_{\alpha,\beta=1}^{p} (\lambda_\alpha^* - \lambda_\alpha)(\lambda_\beta^* - \lambda_\beta)(\underline{x}^{(\alpha)})^T \underline{x}^{(\beta)}$$
$$+ \sum_{\alpha,\beta=1}^{p} (\lambda_\alpha - \lambda_\alpha^*)(\lambda_\beta^* - \lambda_\beta)(\underline{x}^{(\alpha)})^T \underline{x}^{(\beta)}$$
$$- \sum_{\alpha=1}^{p}\lambda_\alpha y_T^{(\alpha)} + \sum_{\alpha=1}^{p}\lambda_\alpha^* y_T^{(\alpha)}$$

$$\boxed{\max_{\lambda_\alpha, \lambda_\alpha^*} \; -\frac{1}{2}\sum_{\alpha,\beta=1}^{p} (\lambda_\alpha^* - \lambda_\alpha)(\lambda_\beta^* - \lambda_\beta)(\underline{x}^{(\alpha)})^T \underline{x}^{(\beta)} + \sum_{\alpha=1}^{p}(\lambda_\alpha^* - \lambda_\alpha) y_T^{(\alpha)}}$$

and we have $\forall\alpha$, $\lambda_\alpha \geq 0$
$\lambda_\alpha^* \geq 0$
$\mu_\alpha \geq 0$
$\mu_\alpha^* \geq 0$
$\delta \geq 0$

$$\mu_\alpha = \frac{C}{p} - \lambda_\alpha \geq 0 \quad\Rightarrow\quad \frac{C}{p} \geq \lambda_\alpha$$

$$\mu_\alpha^* = \frac{C}{p} - \lambda_\alpha^* \geq 0 \quad\Rightarrow\quad \frac{C}{p} \geq \lambda_\alpha^*$$

$$\delta = C\nu - \sum_{\alpha=1}^{p}(\lambda_\alpha + \lambda_\alpha^*) \geq 0 \quad\Rightarrow\quad C\nu \geq \sum_{\alpha=1}^{p}(\lambda_\alpha + \lambda_\alpha^*)$$

so the constraints are:

$$\boxed{\begin{array}{lll} \forall\alpha \in \{1,\dots,p\}: & 0 \leq \lambda_\alpha \leq \frac{C}{p} & 0 \leq \lambda_\alpha^* \leq \frac{C}{p} \\ \sum_{\alpha=1}^{p}(\lambda_\alpha^* - \lambda_\alpha) = 0 & & \sum_{\alpha=1}^{p}(\lambda_\alpha + \lambda_\alpha^*) \leq \nu C \end{array}}$$