The Lagrangian is given by

$$\begin{aligned}
L(w,\epsilon,\xi,b,\eta,\lambda,\mu,\gamma) &= \frac{1}{2}w^\top w - \nu\eta + K_s(\epsilon^\top\epsilon + \xi^\top\xi) + w^\top X\begin{pmatrix}\lambda\\ \mu\end{pmatrix}\\
&\quad - \epsilon^\top\lambda - \xi^\top\mu + b(\mathbf{1}_p^\top\lambda - \mathbf{1}_q^\top\mu) + \eta(\mathbf{1}_p^\top\lambda + \mathbf{1}_q^\top\mu) - \gamma\eta\\
&= \frac{1}{2}w^\top w + w^\top X\begin{pmatrix}\lambda\\ \mu\end{pmatrix} + \eta(\mathbf{1}_p^\top\lambda + \mathbf{1}_q^\top\mu - \nu - \gamma)\\
&\quad + K_s(\epsilon^\top\epsilon + \xi^\top\xi) - \epsilon^\top\lambda - \xi^\top\mu + b(\mathbf{1}_p^\top\lambda - \mathbf{1}_q^\top\mu).
\end{aligned}$$

To find the dual function $G(\lambda,\mu,\gamma)$ we minimize $L(w,\epsilon,\xi,b,\eta,\lambda,\mu,\gamma)$ with respect to w,ϵ,ξ, b, and η. Since the Lagrangian is convex and $(w,\epsilon,\xi,b,\eta)\in\mathbb{R}^n\times\mathbb{R}^p\times\mathbb{R}^q\times\mathbb{R}\times\mathbb{R}$, a convex open set, by Theorem 40.13, the Lagrangian has a minimum in (w,ϵ,ξ,b,η) iff $\nabla L_{w,\epsilon,\xi,b,\eta}=0$, so we compute $\nabla L_{w,\epsilon,\xi,b,\eta}$. The gradient $\nabla L_{w,\epsilon,\xi,b,\eta}$ is given by

$$\nabla L_{w,\epsilon,\xi,b,\eta} = \begin{pmatrix} w + X\begin{pmatrix}\lambda\\ \mu\end{pmatrix}\\ 2K_s\epsilon - \lambda\\ 2K_s\xi - \mu\\ \mathbf{1}_p^\top\lambda - \mathbf{1}_q^\top\mu\\ \mathbf{1}_p^\top\lambda + \mathbf{1}_q^\top\mu - \nu - \gamma\end{pmatrix}.$$

By setting $\nabla L_{w,\epsilon,\xi,b,\eta}=0$ we get the equations

$$\begin{aligned}
w &= -X\begin{pmatrix}\lambda\\ \mu\end{pmatrix}, && (*_w)\\
2K_s\epsilon &= \lambda\\
2K_s\xi &= \mu\\
\mathbf{1}_p^\top\lambda &= \mathbf{1}_q^\top\mu\\
\mathbf{1}_p^\top\lambda + \mathbf{1}_q^\top\mu &= \nu + \gamma.
\end{aligned}$$

The last two equations are identical to the last two equations obtained in Problem $(\mathrm{SVM}_{s2'})$. We can use the other equations to obtain the following expression for the dual function $G(\lambda,\mu,\gamma)$,

$$\begin{aligned}
G(\lambda,\mu,\gamma) &= -\frac{1}{4K_s}(\lambda^\top\lambda + \mu^\top\mu) - \frac{1}{2}\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix}X^\top X\begin{pmatrix}\lambda\\ \mu\end{pmatrix}\\
&= -\frac{1}{2}\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix}\left(X^\top X + \frac{1}{2K_s}I_{p+q}\right)\begin{pmatrix}\lambda\\ \mu\end{pmatrix}.
\end{aligned}$$

Consequently the dual program is equivalent to the minimization program