we get

$$\frac{1}{2}\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix} X^\top X \begin{pmatrix}\lambda \\ \mu\end{pmatrix} - K_m\eta + K_s\left(\sum_{i=1}^p \epsilon_i + \sum_{j=1}^q \xi_j\right) = -\frac{1}{2}\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix} X^\top X \begin{pmatrix}\lambda \\ \mu\end{pmatrix},$$

which yields

$$\eta = \frac{K_s}{K_m}\left(\sum_{i=1}^p \epsilon_i + \sum_{j=1}^q \xi_j\right) + \frac{1}{K_m}\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix} X^\top X \begin{pmatrix}\lambda \\ \mu\end{pmatrix}. \tag{$*$}$$

Therefore, we confirm that $\eta \geq 0$.

Remarks: Since we proved that if the Primal Problem $(\mathrm{SVM}_{s2'})$ has an optimal solution with $w \neq 0$, then $\eta \geq 0$, one might wonder why the constraint $\eta \geq 0$ was included. If we delete this constraint, it is easy to see that the only difference is that instead of the equation

$$\mathbf{1}_p^\top \lambda + \mathbf{1}_q^\top \mu = K_m + \gamma \tag{$*_1$}$$

we obtain the equation

$$\mathbf{1}_p^\top \lambda + \mathbf{1}_q^\top \mu = K_m. \tag{$*_2$}$$

If $\eta > 0$, then by complementary slackness $\gamma = 0$, in which case $(*_1)$ and $(*_2)$ are equivalent. But if $\eta = 0$, then γ could be strictly positive.

The option to omit the constraint $\eta \geq 0$ in the primal is slightly advantageous because then the dual involves $2(p+q)$ instead of $2(p+q)+1$ Lagrange multipliers, so the constraint matrix is a $(p+q+2) \times 2(p+q)$ matrix instead of a $(p+q+2) \times (2(p+q)+1)$ matrix and the matrix defining the quadratic functional is a $2(p+q) \times 2(p+q)$ matrix instead of a $(2(p+q)+1) \times (2(p+q)+1)$ matrix; see Section 54.8.

Under the **Standard Margin Hypothesis** for $(\mathrm{SVM}_{s2'})$, there is some i_0 such that $0 < \lambda_{i_0} < K_s$ and some j_0 such that $0 < \mu_{j_0} < K_s$, and by the complementary slackness conditions $\epsilon_{i_0} = 0$ and $\xi_{j_0} = 0$, so we have the two active constraints

$$w^\top u_{i_0} - b = \eta, \quad -w^\top v_{j_0} + b = \eta,$$

and we can solve for b and η and we get

$$\begin{aligned} b &= \frac{w^\top u_{i_0} + w^\top v_{j_0}}{2} \\ \eta &= \frac{w^\top u_{i_0} - w^\top v_{j_0}}{2} \\ \delta &= \frac{\eta}{\|w\|}. \end{aligned}$$

Due to numerical instability, when writing a computer program it is preferable to compute the lists of indices I_λ and I_μ given by

$$\begin{aligned} I_\lambda &= \{i \in \{1, \ldots, p\} \mid 0 < \lambda_i < K_s\} \\ I_\mu &= \{j \in \{1, \ldots, q\} \mid 0 < \mu_j < K_s\}. \end{aligned}$$