The second and third equations are equivalent to the box constraints

$$0 \leq \lambda_i, \mu_j \leq K_s, \quad i = 1, \ldots, p, \; j = 1, \ldots, q.$$

Since we assumed that the primal problem has an optimal solution with $w \neq 0$, we have

$$X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \neq 0.$$

Plugging back w from $(*_w)$ and b from $(*_b)$ into the Lagrangian, we get

$$\begin{aligned} G(\lambda, \mu, \alpha, \beta) &= \frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} - \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + \frac{1}{2} b^2 - b^2 \\ &= -\frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} - \frac{1}{2} b^2 \\ &= -\frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \left(X^\top X + \begin{pmatrix} \mathbf{1}_p \mathbf{1}_p^\top & -\mathbf{1}_p \mathbf{1}_q^\top \\ -\mathbf{1}_q \mathbf{1}_p^\top & \mathbf{1}_q \mathbf{1}_q^\top \end{pmatrix} \right) \begin{pmatrix} \lambda \\ \mu \end{pmatrix}, \end{aligned}$$

so the dual function is independent of α, β and is given by

$$G(\lambda, \mu) = -\frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \left(X^\top X + \begin{pmatrix} \mathbf{1}_p \mathbf{1}_p^\top & -\mathbf{1}_p \mathbf{1}_q^\top \\ -\mathbf{1}_q \mathbf{1}_p^\top & \mathbf{1}_q \mathbf{1}_q^\top \end{pmatrix} \right) \begin{pmatrix} \lambda \\ \mu \end{pmatrix}.$$

The dual program is given by

$$\begin{aligned} &\text{maximize} \quad -\frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \left(X^\top X + \begin{pmatrix} \mathbf{1}_p \mathbf{1}_p^\top & -\mathbf{1}_p \mathbf{1}_q^\top \\ -\mathbf{1}_q \mathbf{1}_p^\top & \mathbf{1}_q \mathbf{1}_q^\top \end{pmatrix} \right) \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \\ &\text{subject to} \\ &\qquad \sum_{i=1}^p \lambda_i + \sum_{j=1}^q \mu_j = \nu \\ &\qquad 0 \leq \lambda_i \leq K_s, \quad i = 1, \ldots, p \\ &\qquad 0 \leq \mu_j \leq K_s, \quad j = 1, \ldots, q. \end{aligned}$$

Finally, the dual program is equivalent to the following minimization program:

Dual of the Soft margin SVM (SVM_{s3}):

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \left(X^\top X + \begin{pmatrix} \mathbf{1}_p \mathbf{1}_p^\top & -\mathbf{1}_p \mathbf{1}_q^\top \\ -\mathbf{1}_q \mathbf{1}_p^\top & \mathbf{1}_q \mathbf{1}_q^\top \end{pmatrix} \right) \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \\ &\text{subject to} \\ &\qquad \sum_{i=1}^p \lambda_i + \sum_{j=1}^q \mu_j = \nu \\ &\qquad 0 \leq \lambda_i \leq K_s, \quad i = 1, \ldots, p \\ &\qquad 0 \leq \mu_j \leq K_s, \quad j = 1, \ldots, q. \end{aligned}$$