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

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

with

$$P = \begin{pmatrix} XX^\top & -XX^\top \\ -XX^\top & XX^\top \end{pmatrix} = \begin{pmatrix} \mathbf{K} & -\mathbf{K} \\ -\mathbf{K} & \mathbf{K} \end{pmatrix}$$

and

$$q = \begin{pmatrix} y \\ -y \end{pmatrix}.$$

The new dual program is

Dual Program ν-SV Regression Version 2

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix} \left(P + \begin{pmatrix}\mathbf{1}_m\mathbf{1}_m^\top & -\mathbf{1}_m\mathbf{1}_m^\top \\ -\mathbf{1}_m\mathbf{1}_m^\top & \mathbf{1}_m\mathbf{1}_m^\top\end{pmatrix}\right) \begin{pmatrix}\lambda \\ \mu\end{pmatrix} + q^\top \begin{pmatrix}\lambda \\ \mu\end{pmatrix} \\ &\text{subject to} \\ &\qquad \sum_{i=1}^m \lambda_i + \sum_{i=1}^m \mu_i = C\nu \\ &\qquad 0 \leq \lambda_i \leq \frac{C}{m}, \quad 0 \leq \mu_i \leq \frac{C}{m}, \quad i = 1, \ldots, m. \end{aligned}$$

Definition 56.1 and Definition 56.2 are unchanged. We have the following version of Proposition 56.2 showing that p_f, q_f, p_m an q_m have direct influence on the choice of ν.

Proposition 56.7. *(1) Let p_f be the number of points x_i such that $\lambda_i = C/m$, and let q_f be the number of points x_i such that $\mu_i = C/m$. We have $p_f + q_f \leq m\nu$.*

(2) Let p_m be the number of points x_i such that $\lambda_i > 0$, and let q_m be the number of points x_i such that $\mu_i > 0$. We have $p_m + q_m \geq m\nu$.

(3) If $p_f \geq 1$ or $q_f \geq 1$, then $\nu \geq 1/m$.

Proof. (1) Let K_λ and K_μ be the sets of indices corresponding to points failing the margin,

$$\begin{aligned} K_\lambda &= \{i \in \{1, \ldots, m\} \mid \lambda_i = C/m\} \\ K_\mu &= \{i \in \{1, \ldots, m\} \mid \mu_i = C/m\}. \end{aligned}$$