so that the objective function $J(w, \epsilon, \xi, b, \eta)$ is given by

$$J(w, \epsilon, \xi, b, \eta) = \frac{1}{2} w^\top w + (p+q) K_s \left(-\nu\eta + \frac{1}{p+q} \begin{pmatrix} \epsilon^\top & \xi^\top \end{pmatrix} \mathbf{1}_{p+q} \right).$$

Since we obtain an equivalent problem by rescaling by a common positive factor, theoretically it is convenient to normalize K_s as

$$K_s = \frac{1}{p+q},$$

in which case $K_m = \nu$. This method is called the *ν-support vector machine*.

Under the **Standard Margin Hypothesis** for (SVM$_{s2'}$), there is some support vector u_{i_0} of type 1 and some support vector v_{j_0} of type 1, 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

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

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

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

Then b and η are given by the following averaging formulae:

$$\begin{aligned} b &= w^\top \left(\left(\sum_{i \in I_\lambda} u_i \right) / |I_\lambda| + \left(\sum_{j \in I_\mu} v_j \right) / |I_\mu| \right) /2 \\ \eta &= w^\top \left(\left(\sum_{i \in I_\lambda} u_i \right) / |I_\lambda| - \left(\sum_{j \in I_\mu} v_j \right) / |I_\mu| \right) /2. \end{aligned}$$

Proposition 54.1 yields bounds on ν for the method to converge, namely

$$\max \left\{ \frac{2p_f}{p+q}, \frac{2q_f}{p+q} \right\} \leq \nu \leq \min \left\{ \frac{2p_m}{p+q}, \frac{2q_m}{p+q} \right\}.$$