Soft margin SVM $(\mathrm{SVM}_{s2'})$:

$$\begin{aligned}
&\text{minimize} \quad \frac{1}{2}w^\top w - K_m\eta + K_s\left(\epsilon^\top \quad \xi^\top\right)\mathbf{1}_{p+q}\\
&\text{subject to}\\
&\qquad w^\top u_i - b \geq \eta - \epsilon_i, \quad \epsilon_i \geq 0 \qquad i = 1, \ldots, p\\
&\qquad -w^\top v_j + b \geq \eta - \xi_j, \quad \xi_j \geq 0 \qquad j = 1, \ldots, q\\
&\qquad \eta \geq 0.
\end{aligned}$$

This version of the SVM problem was first discussed in Schölkopf, Smola, Williamson, and Bartlett [147] under the name of *ν-SVC* (or *ν-SVM*), and also used in Schölkopf, Platt, Shawe–Taylor, and Smola [146]. The ν-SVC method is also presented in Schölkopf and Smola [145] (which contains much more). The difference between the ν-SVC method and the method presented in Section 54.3, sometimes called the C-SVM method, was thoroughly investigated by Chan and Lin [36].

For this problem it is no longer clear that if $(w, \eta, b, \epsilon, \xi)$ is an optimal solution, then $w \neq 0$ and $\eta > 0$. In fact, if the sets of points are not linearly separable and if K_s is chosen too big, Problem $(\mathrm{SVM}_{s2'})$ may fail to have an optimal solution.

We show that in order for the problem to have a solution we must pick K_m and K_s so that

$$K_m \leq \min\{2pK_s, 2qK_s\}.$$

If we define ν by

$$\nu = \frac{K_m}{(p+q)K_s},$$

then $K_m \leq \min\{2pK_s, 2qK_s\}$ is equivalent to

$$\nu \leq \min\left\{\frac{2p}{p+q}, \frac{2q}{p+q}\right\} \leq 1.$$

The reason for introducing ν is that $\nu(p+q)/2$ can be interpreted as the maximum number of points failing to achieve the margin $\delta = \eta/\|w\|$. We will show later that if the points u_i and v_j are not separable, then we must pick ν so that $\nu \geq 2/(p+q)$ for the method to have a solution for which $w \neq 0$ and $\eta > 0$.

The objective function of our problem is convex and the constraints are affine. Consequently, by Theorem 50.17(2) *if* the Primal Problem $(\mathrm{SVM}_{s2'})$ has an optimal solution, *then* the dual problem has a solution too, and the duality gap is zero. This does not immediately imply that an optimal solution of the dual yields an optimal solution of the primal because the hypotheses of Theorem 50.17(1) fail to hold.

We show that if the primal problem has an optimal solution $(w, \eta, \epsilon, \xi, b)$ with $w \neq 0$, then any optimal solution of the dual problem determines λ and μ, which in turn determine