so η is also expressed in terms of λ, μ.

The condition $\nu > (p_f + q_f)/(p+q)$ cannot be satisfied if $p_f + q_f = p+q$, but in this case all points fail the margin, which indicates that δ is too big, so we reduce ν and try again.

Remark: The equation

$$\sum_{i=1}^{p} \lambda_i + \sum_{j=1}^{q} \mu_j = \nu$$

implies that either there is some i_0 such that $\lambda_{i_0} > 0$ *or* there is some j_0 such that $\mu_{j_0} > 0$, which implies that $p_m + q_m \geq 1$.

Another way to compute η is to assume the Standard Margin Hypothesis for (SVM_{s3}). Under the **Standard Margin Hypothesis** for (SVM_{s3}), either there is some i_0 such that $0 < \lambda_{i_0} < K_s$ *or* there is some j_0 such that $0 < \mu_{j_0} < K_s$, in other words, there is some support vector of type 1. By the complementary slackness conditions $\epsilon_{i_0} = 0$ or $\xi_{j_0} = 0$, so we have

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

and we can solve for η.

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

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

Then it is easy to see that we can compute η using the following averaging formulae: If $I_\lambda \neq \emptyset$, then

$$\eta = w^\top \left(\sum_{i \in I_\lambda} u_i \right) / |I_\lambda| - b,$$

and if $I_\mu \neq \emptyset$, then

$$\eta = b - w^\top \left(\sum_{j \in I_\mu} v_j \right) / |I_\mu|.$$

Theoretically the condition $\nu > (p_f + q_f)/(p+q)$ is less restrictive that the **Standard Margin Hypothesis** but in practice we have never observed an example for which $\nu > (p_f + q_f)/(p+q)$ and yet the **Standard Margin Hypothesis** fails.

The "kernelized" version of Problem (SVM_{s3}) is the following:

Soft margin kernel SVM (SVM_{s3}):

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