so we get

$$w = \frac{\displaystyle\sum_{i=1}^{p} \lambda_i u_i - \sum_{j=1}^{q} \mu_j v_j}{\left(\begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix}\right)^{1/2}},$$

which is the result of making $\sum_{i=1}^{p} \lambda_i u_i - \sum_{j=1}^{q} \mu_j v_j$ a unit vector, since

$$X = \begin{pmatrix} -u_1 & \cdots & -u_p & v_1 & \cdots & v_q \end{pmatrix}.$$

It remains to find b and δ, which are not given by the dual program and for this we use the complementary slackness conditions.

The equations

$$\sum_{i=1}^{p} \lambda_i = \sum_{j=1}^{q} \mu_j = \frac{1}{2}$$

imply that there is some i_0 such that $\lambda_{i_0} > 0$ and some j_0 such that $\mu_{j_0} > 0$, but a priori, nothing prevents the situation where $\lambda_i = K$ for all nonzero λ_i or $\mu_j = K$ for all nonzero μ_j. If this happens, we can rerun the optimization method with a larger value of K. If the following mild hypothesis holds, then b and δ can be found.

Standard Margin Hypothesis for (SVM$_{s1}$). There is some index i_0 such that $0 < \lambda_{i_0} < K$ and there is some index j_0 such that $0 < \mu_{j_0} < K$. This means that some u_{i_0} is a support vector of type 1 on the blue margin, and some v_{j_0} is a support of type 1 on the red margin.

If the **Standard Margin Hypothesis** for (SVM$_{s1}$) holds, then $\epsilon_{i_0} = 0$ and $\mu_{j_0} = 0$, and then we have the active equations

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

and we obtain the values of b and δ as

$$\begin{aligned} b &= \frac{1}{2}(w^\top u_{i_0} + w^\top v_{j_0}) \\ \delta &= \frac{1}{2}(w^\top u_{i_0} - w^\top v_{j_0}). \end{aligned}$$

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

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