This time we obtain w, b, η, ϵ and ξ from λ and μ:

$$\begin{aligned} w &= \sum_{i=1}^{p} \lambda_i u_i - \sum_{j=1}^{q} \mu_j v_j \\ b &= -\sum_{i=1}^{p} \lambda_i + \sum_{j=1}^{q} \mu_j \\ \epsilon &= \frac{\lambda}{2K} \\ \xi &= \frac{\mu}{2K}, \end{aligned}$$

and

$$(p+q)K_s\nu\eta = \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \left(X^\top X + \begin{pmatrix} \mathbf{1}_p\mathbf{1}_p^\top & -\mathbf{1}_p\mathbf{1}_q^\top \\ -\mathbf{1}_q\mathbf{1}_p^\top & \mathbf{1}_q\mathbf{1}_q^\top \end{pmatrix} + \frac{1}{2K_s} I_{p+q} \right) \begin{pmatrix} \lambda \\ \mu \end{pmatrix}.$$

The constraint

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

implies that either there is some i_0 such that $\lambda_{i_0} > 0$ or there is some j_0 such that $\mu_{j_0} > 0$, we have $\epsilon_{i_0} > 0$ or $\xi_{j_0} > 0$, which means that at least one point is misclassified, so Problem (SVM_{s5}) should only be used when the sets $\{u_i\}$ and $\{v_j\}$ are *not* linearly separable.

These methods all have a kernelized version.

We implemented all these methods in `Matlab`, solving the dual using ADMM.

From a theoretical point of view, Problems (SVM_{s4}) and (SVM_{s5}) seem to have more advantages than the others since they determine w, b, η and b without requiring any condition about support vectors of type 1. However, from a practical point of view, Problems (SVM_{s4}) and (SVM_{s5}) are less flexible that ($\mathrm{SVM}_{s2'}$) and (SVM_{s3}), and we have observed that (SVM_{s4}) and (SVM_{s5}) are unable to produce as small a margin δ as ($\mathrm{SVM}_{s2'}$) and (SVM_{s3}).

54.18 Problems

Problem 54.1. Prove the following inequality

$$\max\left\{\frac{1}{2p_m}, \frac{1}{2q_m}\right\} \le K \le \min\left\{\frac{1}{2p_f}, \frac{1}{2q_f}\right\}$$

stated just after Definition 54.1.