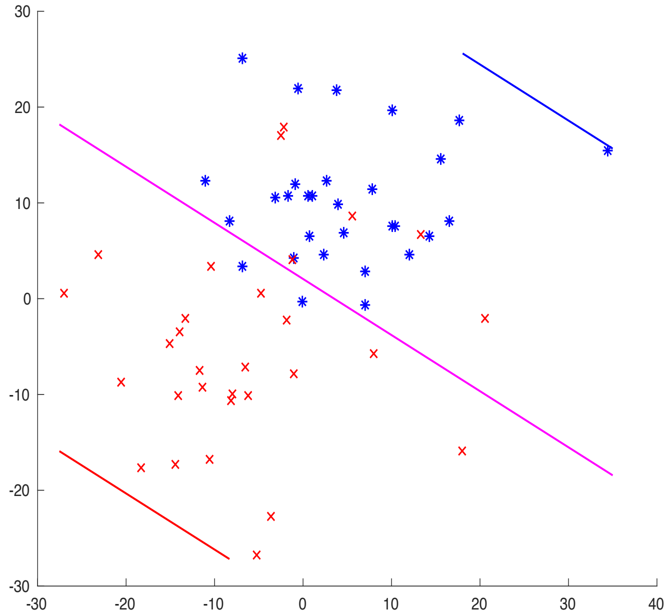

Figure 54.23: Running (SVM$_{s4}$) on two sets of 30 points; $K = 1/12000$.

method *does not* require support vectors to compute b. We can omit the constraint $\eta \geq 0$, because for an optimal solution it can be shown using duality that $\eta \geq 0$.

A drawback of Program (SVM$_{s5}$) is that for fixed K_s, the quantity $\delta = \eta/\|w\|$ and the hyperplanes $H_{w,b}, H_{w,b+\eta}$ and $H_{w,b-\eta}$ are independent of ν. This will be shown in Theorem 54.9. Thus this method is less flexible than (SVM$_{s2'}$) and (SVM$_{s3}$).

The Lagrangian is given by

$$\begin{aligned}
L(w, \epsilon, \xi, b, \eta, \lambda, \mu) &= \frac{1}{2}w^\top w + \frac{1}{2}b^2 - \nu\eta + K_s(\epsilon^\top\epsilon + \xi^\top\xi) + w^\top X \begin{pmatrix}\lambda \\ \mu\end{pmatrix} \\
&\quad - \epsilon^\top\lambda - \xi^\top\mu + b(\mathbf{1}_p^\top\lambda - \mathbf{1}_q^\top\mu) + \eta(\mathbf{1}_p^\top\lambda + \mathbf{1}_q^\top\mu) \\
&= \frac{1}{2}w^\top w + w^\top X \begin{pmatrix}\lambda \\ \mu\end{pmatrix} + \eta(\mathbf{1}_p^\top\lambda + \mathbf{1}_q^\top\mu - \nu) \\
&\quad + K_s(\epsilon^\top\epsilon + \xi^\top\xi) - \epsilon^\top\lambda - \xi^\top\mu + b(\mathbf{1}_p^\top\lambda - \mathbf{1}_q^\top\mu) + \frac{1}{2}b^2.
\end{aligned}$$

To find the dual function $G(\lambda, \mu)$ we minimize $L(w, \epsilon, \xi, b, \eta, \lambda, \mu)$ with respect to w, ϵ, ξ, b, and η. Since the Lagrangian is convex and $(w, \epsilon, \xi, b, \eta) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}$, a convex open set, by Theorem 40.13, the Lagrangian has a minimum in $(w, \epsilon, \xi, b, \eta)$ iff $\nabla L_{w,\epsilon,\xi,b,\eta} = 0$, so