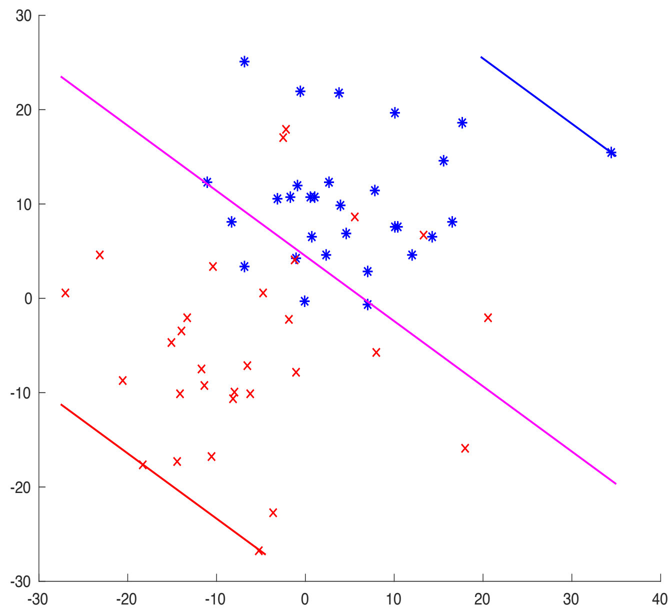

Figure 54.15: Running $(\mathrm{SVM}_{s2'})$ on two sets of 30 points; $\nu = 0.97$.

To find the dual function $G(\lambda, \mu, \alpha, \beta)$, we minimize $L(w, \epsilon, \xi, b, \eta, \lambda, \mu, \alpha, \beta)$ 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 we compute its gradient with respect to $w, \epsilon, \xi, b, \eta$, and we get

$$\nabla L_{w,\epsilon,\xi,b,\eta} = \begin{pmatrix} X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + w \\ K_s \mathbf{1}_p - (\lambda + \alpha) \\ K_s \mathbf{1}_q - (\mu + \beta) \\ b + \mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu \\ \mathbf{1}_p^\top \lambda + \mathbf{1}_q^\top \mu - \nu \end{pmatrix}.$$

By setting $\nabla L_{w,\epsilon,\xi,b,\eta} = 0$ we get the equations

$$w = -X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \qquad (*_w)$$

$$\begin{aligned} \lambda + \alpha &= K_s \mathbf{1}_p \\ \mu + \beta &= K_s \mathbf{1}_q \\ \mathbf{1}_p^\top \lambda + \mathbf{1}_q^\top \mu &= \nu, \end{aligned}$$

and

$$b = -(\mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu). \qquad (*_b)$$