Dual Program ϵ-SV Regression:

$$
\begin{aligned}
&\text{minimize} \quad \frac{1}{2}\sum_{i,j=1}^{m}(\lambda_i-\mu_i)(\lambda_j-\mu_j)x_i^\top x_j+\sum_{i=1}^{m}(\lambda_i-\mu_i)y_i+\epsilon\sum_{i=1}^{m}(\lambda_i+\mu_i)\\
&\text{subject to}\\
&\qquad \sum_{i=1}^{m}\lambda_i-\sum_{i=1}^{m}\mu_i=0\\
&\qquad 0\leq\lambda_i\leq\frac{C}{m},\quad 0\leq\mu_i\leq\frac{C}{m},\quad i=1,\ldots,m,
\end{aligned}
$$

minimizing over α and μ.

The constraint

$$\sum_{i=1}^{m}\lambda_i+\sum_{i=1}^{m}\mu_i\leq C\nu$$

is gone but the extra term $\epsilon\sum_{i=1}^{m}(\lambda_i+\mu_i)$ has been added to the dual function, to prevent λ_i and μ_i from blowing up.

There is an obvious kernelized version of ϵ-SV regression. It is easy to show that ν-SV regression subsumes ϵ-SV regression, in the sense that if ν-SV regression succeeds and yields $w, b, \epsilon > 0$, then ϵ-SV regression with the same C and the same value of ϵ also succeeds and returns the same pair (w, b). For more details on these methods, see Schölkopf, Smola, Williamson, and Bartlett [147].

Remark: The linear penalty function $\sum_{i=1}^{m}(\xi_i+\xi'_i)$ can be replaced by the quadratic penalty function $\sum_{i=1}^{m}(\xi_i^2+\xi_i'^2)$; see Shawe–Taylor and Christianini [159] (Chapter 7). In this case, it is easy to see that for an optimal solution we must have $\xi_i \geq 0$ and $\xi'_i \geq 0$, so we may omit the constraints $\xi_i \geq 0$ and $\xi'_i \geq 0$. We must also have $\gamma = 0$ so we omit the variable γ as well. It can be shown that $\xi = (m/2C)\lambda$ and $\xi' = (m/2C)\mu$. This problem is very similar to the Soft Margin SVM (SVM_{s4}) discussed in Section 54.13.

56.5 ν-Regression Version 2; Penalizing b

Yet another variant of ν-SV regression is to add the term $\frac{1}{2}b^2$ to the objective function. We will see that solving the dual not only determines w but also b and ϵ (provided a mild condition on ν). We wish to solve the following program: