As in the previous section, we assume that our data points $\{x_1, \ldots, x_m\}$ belong to a set $\mathcal{X}$ and we pretend that we have feature space $(F, \langle -, - \rangle)$ and a feature embedding map $\varphi\colon \mathcal{X} \to F$, but we only have access to the kernel function $\kappa(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle$. We wish to perform ν-SV regression in the feature space F on the data set $\{(\varphi(x_1), y_1), \ldots, (\varphi(x_m), y_m)\}$. Going over the previous computation, we see that the primal program is given by

Program kernel ν-SV Regression:

$$
\begin{aligned}
&\text{minimize} \quad \frac{1}{2}\langle w, w\rangle + C\left(\nu\epsilon + \frac{1}{m}\sum_{i=1}^{m}(\xi_i + \xi_i')\right)\\
&\text{subject to}\\
&\qquad \langle w, \varphi(x_i)\rangle + b - y_i \leq \epsilon + \xi_i, \quad \xi_i \geq 0 \qquad i = 1, \ldots, m\\
&\qquad -\langle w, \varphi(x_i)\rangle - b + y_i \leq \epsilon + \xi_i', \quad \xi_i' \geq 0 \qquad i = 1, \ldots, m\\
&\qquad \epsilon \geq 0,
\end{aligned}
$$

minimizing over the variables w, ϵ, b, ξ, and ξ'.

The Lagrangian is given by

$$
\begin{aligned}
L(w, b, \lambda, \mu, \gamma, \xi, \xi', \epsilon, \alpha, \beta) &= \frac{1}{2}\langle w, w\rangle + \left\langle w, \sum_{i=1}^{m}(\lambda_i - \mu_i)\varphi(x_i)\right\rangle\\
&\quad + \epsilon\left(C\nu - \gamma - \sum_{i=1}^{m}(\lambda_i + \mu_i)\right) + \sum_{i=1}^{m}\xi_i\left(\frac{C}{m} - \lambda_i - \alpha_i\right)\\
&\quad + \sum_{i=1}^{m}\xi_i'\left(\frac{C}{m} - \mu_i - \beta_i\right) + b\left(\sum_{i=1}^{m}(\lambda_i - \mu_i)\right) - \sum_{i=1}^{m}(\lambda_i - \mu_i)y_i.
\end{aligned}
$$

Setting the gradient $\nabla L_{w,\epsilon,b,\xi,\xi'}$ of the Lagrangian to zero, we also obtain the equations

$$
w = \sum_{i=1}^{m}(\mu_i - \lambda_i)\varphi(x_i), \tag{$*_w$}
$$

$$
\begin{aligned}
\sum_{i=1}^{m}\lambda_i - \sum_{i=1}^{m}\mu_i &= 0\\
\sum_{i=1}^{m}\lambda_i + \sum_{i=1}^{m}\mu_i + \gamma &= C\nu\\
\lambda + \alpha = \frac{C}{m}, \quad \mu + \beta &= \frac{C}{m}.
\end{aligned}
$$

Using the above equations, we find that the dual function G is independent of the variables β, α, β, and we obtain the following dual program: