**Program (RR3$b$):**

$$\begin{aligned}
&\text{minimize} \quad \xi^\top \xi + K w^\top w + K b^2 \\
&\text{subject to} \\
&\qquad y - [X\ \mathbf{1}] \begin{pmatrix} w \\ b \end{pmatrix} = \xi,
\end{aligned}$$

minimizing over $\xi, w$ and $b$.

This program is solved just as Program (**RR2**). In terms of the dual variable $\alpha$, we get

$$\begin{aligned}
\alpha &= ([X\ \mathbf{1}][X\ \mathbf{1}]^\top + K I_m)^{-1} y \\
\begin{pmatrix} w \\ b \end{pmatrix} &= [X\ \mathbf{1}]^\top \alpha \\
\xi &= K\alpha.
\end{aligned}$$

Thus $b = \mathbf{1}^\top \alpha$. Observe that $[X\ \mathbf{1}][X\ \mathbf{1}]^\top = XX^\top + \mathbf{1}\mathbf{1}^\top$.

If $n < m$, it is preferable to use the formula

$$\begin{pmatrix} w \\ b \end{pmatrix} = ([X\ \mathbf{1}]^\top [X\ \mathbf{1}] + K I_{n+1})^{-1} [X\ \mathbf{1}]^\top y.$$

Since we also have the equation

$$y - Xw - b\mathbf{1} = \xi,$$

we obtain

$$\frac{1}{m}\mathbf{1}^\top y - \frac{1}{m}\mathbf{1}^\top X w - \frac{1}{m} b \mathbf{1}^\top \mathbf{1} = \frac{1}{m}\mathbf{1}^\top K\alpha,$$

so

$$\overline{y} - (\overline{X^1}\ \cdots\ \overline{X^n}) w - b = \frac{1}{m} K b,$$

which yields

$$b = \frac{m}{m+K}(\overline{y} - (\overline{X^1}\ \cdots\ \overline{X^n}) w).$$

**Remark:** As a least squares problem, the solution is given in terms of the pseudo-inverse $[X\ \mathbf{1}]^+$ of $[X\ \mathbf{1}]$ by

$$\begin{pmatrix} w \\ b \end{pmatrix} = [X\ \mathbf{1}]^+ y.$$

**Example 55.2.** Applying Program (**RR3$b$**) to the data set of Example 55.1 with $K = 0.01$ yields

$$w = \begin{pmatrix} 1.1706 \\ 1.1401 \end{pmatrix}, \quad b = -1.2298.$$