we get

$$\mu = (\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1})^{-1}\mathbf{1}^\top (XX^\top + KI_m)^{-1} y, \quad \alpha = (XX^\top + KI_m)^{-1}(y - \mu\mathbf{1}).$$

Note that the matrix $B^\top AB$ is the scalar $\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1}$, which is the negative of the Schur complement of $XX^\top + KI_m$.

Interestingly $b = \mu$, which is not obvious a priori.

Proposition 55.2. *We have $b = \mu$.*

Proof. To prove this result we need to express α differently. Since μ is a scalar, $\mu\mathbf{1} = \mathbf{1}\mu$, so

$$\mu\mathbf{1} = \mathbf{1}\mu = (\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{1}^\top (XX^\top + KI_m)^{-1} y,$$

and we obtain

$$\alpha = (XX^\top + KI_m)^{-1}(I_m - (\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{1}^\top (XX^\top + KI_m)^{-1})y. \tag{$*_{\alpha_3}$}$$

Since $w = X^\top\alpha$, we have

$$w = X^\top (XX^\top + KI_m)^{-1}(I_m - (\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{1}^\top (XX^\top + KI_m)^{-1})y. \tag{$*_{w_3}$}$$

From $\xi = K\alpha$, we deduce that b is given by the equation

$$b\mathbf{1} = y - Xw - K\alpha.$$

Since $w = X^\top\alpha$, using $(*_{\alpha_3})$ we obtain

$$\begin{aligned} b\mathbf{1} &= y - Xw - K\alpha \\ &= y - (XX^\top + KI_m)\alpha \\ &= y - (I_m - (\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{1}^\top (XX^\top + KI_m)^{-1})y \\ &= (\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{1}^\top (XX^\top + KI_m)^{-1})y \\ &= \mu\mathbf{1}, \end{aligned}$$

and thus

$$b = \mu = (\mathbf{1}^\top (XX^\top + KI_m)^{-1}\mathbf{1})^{-1}\mathbf{1}^\top (XX^\top + KI_m)^{-1} y, \tag{$*_{b_3}$}$$

as claimed. ☐

In summary the KKT-equations determine both α and μ, and so $w = X^\top\alpha$ and b as well.

There is also a useful expression of b as an average.

Since $\mathbf{1}^\top\mathbf{1} = m$ and $\mathbf{1}^\top\alpha = 0$, we get

$$b = \frac{1}{m}\mathbf{1}^\top y - \frac{1}{m}\mathbf{1}^\top Xw - \frac{1}{m}K\mathbf{1}^\top\alpha = \overline{y} - \sum_{j=1}^{n} \overline{X^j} w_j,$$