Program (Dlasso3):

$$\begin{aligned} &\text{minimize} && \frac{1}{2}\|y-\lambda\|_2^2 \\ &\text{subject to} && \\ & && \left\|X^\top\lambda\right\|_\infty \leq \tau \\ & && \mathbf{1}_m^\top\lambda = 0, \end{aligned}$$

minimizing over $\lambda \in \mathbb{R}^m$.

Once $\lambda = \xi$ and w are determined, we obtain b using the equation

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

and since $\mathbf{1}_m^\top\mathbf{1}_m = m$ and $\mathbf{1}_m^\top\xi = \mathbf{1}_m^\top\lambda = 0$, the above yields

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

where $\overline{y}$ is the mean of y and $\overline{X^j}$ is the mean of the jth column of X.

The equation

$$b = \widehat{b} + \overline{y} - \sum_{j=1}^{n}\overline{X^j}w_j = \widehat{b} + \overline{y} - (\overline{X^1}\ \cdots\ \overline{X^n})w,$$

can be used as in ridge regression, (see Section 55.2), to show that the Program (**lasso3**) is equivalent to applying lasso regression (**lasso2**) without an intercept term to the centered data, by replacing y by $\widehat{y} = y - \overline{y}\mathbf{1}$ and X by $\widehat{X} = X - \overline{X}$. Then b is given by

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

where $\widehat{w}$ is the solution given by (**lasso2**). This is the method described by Hastie, Tibshirani, and Wainwright [89] (Section 2.2).

Example 55.3. We can create a data set (X, y) where X a 100×5 matrix and y is a 100×1 vector using the following `Matlab` program in which the command `randn` creates an array of normally distributed numbers.

```
X = randn(100,5);
ww = [0; 2; 0; -3; 0];
y = X*ww + randn(100,1)*0.1;
```

The purpose of the third line is to add some small noise to the "output" $X * ww$. The first five rows of X are

$$\begin{pmatrix} -1.1658 & -0.0679 & -1.6118 & 0.3199 & 0.4400 \\ -1.1480 & -0.1952 & -0.0245 & -0.5583 & -0.6169 \\ 0.1049 & -0.2176 & -1.9488 & -0.3114 & 0.2748 \\ 0.7223 & -0.3031 & 1.0205 & -0.5700 & 0.6011 \\ 2.5855 & 0.0230 & 0.8617 & -1.0257 & 0.0923 \end{pmatrix},$$