For $\tau = 0.02$, we have

$$w = \begin{pmatrix} 0.00003 \\ 2.01056 \\ -0.00004 \\ -2.99821 \\ 0.00000 \end{pmatrix}, \quad b = 0.00135.$$

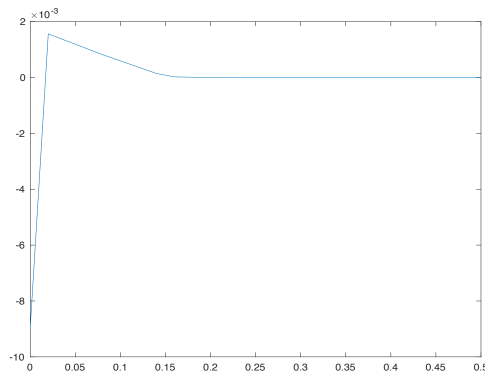

Figure 55.9: Fifth component of w.

This weight vector w is very close to the original vector $ww = [0; 2; 0; -3; 0]$ that we used to create y. For large values of τ, the weight vector is essentially the zero vector. This happens for $\tau = 235$, where every component of w is less than 10^{-5}.

Another way to find b is to add the term $(C/2)b^2$ to the objective function, for some positive constant C, obtaining the program

Program(lasso4):

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}\xi^\top \xi + \tau \mathbf{1}_n^\top \epsilon + \frac{1}{2}Cb^2 \\ &\text{subject to} \\ &\qquad\qquad y - Xw - b\mathbf{1}_m = \xi \\ &\qquad\qquad w \leq \epsilon \\ &\qquad\qquad -w \leq \epsilon, \end{aligned}$$

minimizing over ξ, w, ϵ and b.

This time the Lagrangian is

$$\begin{aligned} L(\xi, w, \epsilon, b, \lambda, \alpha_+, \alpha_-) = {} & \frac{1}{2}\xi^\top \xi - \xi^\top \lambda + \lambda^\top y + \frac{C}{2}b^2 - b\mathbf{1}_m^\top \lambda \\ & + \epsilon^\top(\tau \mathbf{1}_n - \alpha_+ - \alpha_-) + w^\top(\alpha_+ - \alpha_- - X^\top \lambda), \end{aligned}$$