Variable	Units	Mean	Std dev
Median income, x_1	\$	50000	15000
Median age, x_2	years	45	10
House sale price, y	\$1000	300	100

Table 1: Features for Problem 2

1. *Selecting a regularizer.* Suppose we fit a regularized least squares objective,

$$J(\mathbf{w}) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + \lambda \phi(\mathbf{w}),$$

where $\hat{y}_i$ is some prediction of y_i given the model parameters $\mathbf{w}$. For each case below, suggest a possible regularization function $\phi(\mathbf{w})$. There is no single correct answer.

   (a) All parameters vectors $\mathbf{w}$ should be considered.

   (b) Negative values of w_j are unlikely (but still possible).

   (c) For each j, w_j should not change that significantly from w_{j-1}.

   (d) For most j, $w_j = w_{j-1}$. However, it can happen that w_j can be different from w_{j-1} for a few indices j.

2. *Normalization.* A data analyst for a real estate firm wants to predict house prices based on two features in each zip code. The features are shown in Table 1. The agent decides to use a linear model,

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2, \tag{1}$$

   (a) What is the problem in using a LASSO regularizer of the form,

$$\phi(\boldsymbol{\beta}) = \sum_{j=1}^{2} |\beta_j|.$$

   (b) To uniformly regularize the features, she fits a model on the normalized features,

$$\hat{u} = \alpha_1 z_1 + \alpha_2 z_2, \quad z_j = \frac{x_j - \bar{x}_j}{s_j}, \quad u = \frac{\hat{y} - \bar{y}}{s_y},$$

   where s_j and s_y are the standard deviations of the x_j and y. She obtains parameters $\boldsymbol{\alpha} = [0.6, -0.3]$? What are the parameters β in the original model (1)?

3. *Minimizing an ℓ_1 objective.* In this problem, we will show how to minimize a simple scalar function with an ℓ_1-term. Given y and $\lambda > 0$, suppose we wish to find the minimum,

$$\widehat{w} = \arg\min_{w} J(w) = \frac{1}{2}(y - w)^2 + \lambda |w|.$$

Write $\widehat{w}$ in terms of y and λ. Since $|w|$ is not differentiable everywhere, you cannot simple set $J'(w) = 0$ and solve for w. Instead, you have to look at three cases:

(i) First, suppose there is a minima at $w > 0$. In this region, $|w| = w$. Since the set $w > 0$ is open, at any minima $J'(w) = 0$. Solve for w and test if the solution indeed satisfies $w > 0$.

(ii) Similarly, suppose $w < 0$. Solve for $J'(w) = 0$ and test if the solution satisfies the assumption that $w < 0$.

(iii) If neither of the above cases have a minima, then the minima must be at $w = 0$.