Derivative Free Optimization

Exercise 1: Coordinate Descent I
Minimize Ridge regression, i.e.,

$$\min_{\boldsymbol{\theta}} \frac{1}{2}\|\mathbf{X}\boldsymbol{\theta} - \mathbf{y}\|_2^2 + \frac{\lambda}{2}\|\boldsymbol{\theta}\|_2^2$$

for $\lambda \geq 0$ via coordinate descent under the assumption that $\mathbf{X}^\top\mathbf{X} = \mathbf{I}_d$.

Exercise 2: Coordinate Descent II
Consider the function

$$g : \mathbb{R}^2 \to \mathbb{R}, (x_1, x_2) \mapsto |x_1 - x_2| + 0.1(x_1 + x_2).$$

(a) Perform one round of coordinate descent starting from an arbitrary point (x_1, x_2). Show that after updating x_1 (while fixing x_2) and then updating x_2 (while fixing x_1) the algorithm arrives at a point where $x_1 = x_2$ and terminates. That is, show that coordinate descent will not move beyond the first iteration.

(b) Show that the global infimum of g is $-\infty$. Conclude that coordinate descent fails to find the true minimizer for this function.