Machine Learning, 2024 Spring
Assignment 6

Notice

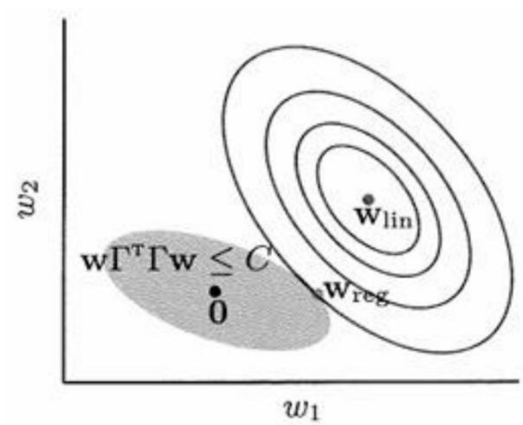
Plagiarizer will get 0 points.
LaTeXis highly recommended. Otherwise you should write as legibly as possible.

Problem 1 In this problem, you will investigate the relationship between the soft order constraint and the augmented error. The regularized weight $\mathbf{w}_{\text{reg}}$ is a solution to

$$\min E_{\text{in}}(\mathbf{w}) \text{ subject to } \mathbf{w}^{\mathrm{T}} \Gamma^{\mathrm{T}} \Gamma \mathrm{w} \leq C$$

(a) If $\mathbf{w}_{\text{lin}}^{\mathrm{T}} \Gamma^{\mathrm{T}} \Gamma \mathrm{w}_{\text{lin}} \leq C$, then what is w_{reg} ?

(b) If $\mathbf{w}_{\text{lin}}^{\tau} \Gamma^{\top} \Gamma \mathbf{w}_{\text{lin}} > C$, the situation is illustrated below,



The constraint is satisfied in the shaded region and the contours of constant E_{in} are the ellipsoids (why ellipsoids?). What is $\mathbf{w}_{\text{reg}}^{\mathrm{T}} \Gamma^{\mathrm{T}} \Gamma \mathbf{w}_{\text{reg}}$?

(c) Show that with

$$\lambda_C = -\frac{1}{2C} \mathbf{w}_{\text{reg}}^{\mathrm{T}} \nabla E_{\text{in}}(\mathbf{w}_{\text{reg}})$$

$\mathbf{w}_{\text{reg}}$ minimizes $E_{\text{in}}(\mathbf{w}) + \lambda_C \mathbf{w}^{\mathrm{T}} \Gamma^{\mathrm{T}} \Gamma \mathbf{w}$. [Hint: use the previous part to solve for w_{reg} as an equality constrained optimization problem using the method of Lagrange multipliers.]

(d) Show that the following hold for λ_C :

(i) If $\mathrm{w}_{\text{lin}}^{\mathrm{T}} \Gamma^{\top} \Gamma \mathrm{w}_{\text{lin}} \leq C$ then $\lambda_C = 0$ (w_{lin} itself satisfies the constraint).

(ii) If $\mathrm{w}_{\text{lin}}^{\mathrm{T}} \Gamma^{\mathrm{T}} \Gamma \mathrm{w}_{\text{lin}} > C$, then $\lambda_C > 0$ (the penalty term is positive).

(iii) If $\mathrm{w}_{\mathrm{lin}}^{\mathrm{T}} \Gamma^{\mathrm{T}} \Gamma \mathrm{w}_{\mathrm{lin}} > C$, then λ_C is a strictly decreasing function of C. [Hint: show that $\frac{d\lambda_C}{dC} < 0$ for $C \in \left[0, \mathbf{w}_{\mathrm{lin}}^{\mathrm{T}} \Gamma^{T} \Gamma \mathbf{w}_{\mathrm{lin}}\right]$].

Problem 2 [The Lasso algorithm] Rather than a soft order constraint on the squares of the weights, one could use the absolute values of the weights:

$$\min E_{\text{in}}(\mathbf{w}) \text{ subject to } \sum_{i=0}^{d} |w_i| \leq C$$

The model is called the lasso algorithm.

(a) Formulate and implement this as a quadratic program.

(b) What is the augmented error and discuss the algorithm for solving it. You can solve this problem using iterative soft-thresholding algorithm or a gradient projection method and present your pseudo-code.

Problem 3 Similar to problem 3 in assignment 3 and assignment 4, you need to use the SUV dataset to implement (using Python or MATLAB) the L1/L2 regularization (penalty/augmented).

- Present your code.
- Present the path plots of L1 and L2 regularization. (Notice: you need to mark the selected value of the regular parameter)
- Analyze the weight difference between L1 and L2 regularization. (Notice: you need to describe the similarities and differences between the solutions of path plots)
- If you only want to build a model that contains 2 variables, which two features would you choose?

Solution: