Please write a report for this project.

Consider training the following regularized logistic regression model

$$\min_{\boldsymbol{x}} F(\boldsymbol{x}) := f(\boldsymbol{x}) + \lambda R(\boldsymbol{x}),$$

where

$$f(\boldsymbol{x}) = \frac{1}{2n} \sum_{i=1}^{n} \log\left(1 + \exp(-b_i \boldsymbol{a}_i^\top \boldsymbol{x})\right),$$

with n being the sample size and $\boldsymbol{a}_i \in \mathbb{R}^d$ $(d = 50)$ is a training data, $b_i \in \{-1, 1\}$ be the label of $\boldsymbol{a}_i$. Here, we consider two different regularization functions i.e. ℓ_1-regularization $(R(\boldsymbol{x}) = \|\boldsymbol{x}\|_1)$ and ℓ_2-regularization $(R(\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2)$.

Please use the code in the zip file to generate 1000 data-label pairs $\{\boldsymbol{a}_i, b_i\}_{i=1}^{1000}$.

- Derive $prox_{\lambda\|\boldsymbol{x}\|_1}(\boldsymbol{x})$ and $prox_{\lambda\|\boldsymbol{x}\|_2}(\boldsymbol{x})$.
- For $\lambda = 0.001$, numerically solve the problem $\min_{\boldsymbol{x}} F(\boldsymbol{x})$ using subgradient method, proximal gradient method, accelerated proximal gradient method with heavy-ball momentum and Nesterov's acceleration. Plot $F(\boldsymbol{x}^k) - F(\boldsymbol{x}^*)$ over the iteration k for each method, where $\boldsymbol{x}^*$ is in the code that used to generate the training data.
- Test different λ, e.g. $0.005, 0.01, 0.05, 0.1$ and see how $\boldsymbol{x}^k$ changes after you run enough number of iterations.
- Can you propose any approach to further accelerate the training process?