

Experiments 4. Given pairs of vectors $(z^i, y^i) \in \mathbb{R}^n \times \{-1, 1\}$, $i = 1, \dots, m$, let us consider the logistic regression function with ℓ_2 -regularization

$$f(x) = \frac{\sigma}{2} \|x\|_2^2 + \sum_{i=1}^m \log(1 + e^{-(x^\top z^i) y^i}), \quad (21)$$

where $\sigma \geq 0$ is a parameter. By straightforward calculations, we obtain

$$\nabla f(x) = \sigma x - \sum_{i=1}^m \frac{y^i h_i(x)}{1 + h_i(x)} z^i$$

and

$$\nabla^2 f(x) = \sigma I + \sum_{i=1}^m \frac{(y^i)^2 h_i(x)}{1 + h_i(x)} \left[1 - \frac{1}{1 + h_i(x)} \right] z^i (z^i)^\top,$$

where $h_i(x) = e^{-(x^\top z^i) y^i}$. Immediately, f is σ -strongly convex if $\sigma > 0$. However, we also consider in our tests the case $\sigma = 0$. Actually, note that $h_i(x) > 0$, $i = 1, \dots, m$, remain bounded when minimizing f , and thus the Rayleigh quotients $(u^\top \nabla^2 f(x) u) / (u^\top u)$ remains uniformly above a positive scalar during the minimization process, at least when z forms a basis for \mathbb{R}^n (usually, $m > n$). That is, $\nabla^2 f(x)$ has great chances to be definite positive with an uniform constant over the sequences generate by a method even if $\sigma = 0$.

The function (21) appears in *binary classification problems*. In fact, note that minimizing the sum in (21) leads each weighted data $x^\top z^i$ to have the same sign as y^i . In this sense, let us consider (21) construct from the Ionosphere dataset, available from the UCI Machine Learning Repository [18]. This dataset consists of 351 radar returns z^i from the ionosphere together a binary label y^i that indicates weather or not each return is good for analysis (in our case, $y = 1$ for good returns and -1 otherwise). Each entry z^i encodes 34 continuous attributes, all normalized to $[-1, 1]$, so $z^i \in [-1, 1]^{34}$ for all i . It is worth be mentioned that we are not training a model/neural network to predict the correct answer to an unknown data, as originally proposed [17]. In particular, we do not divide the dataset into training and test data.