Constraint relaxation algorithm

Definitions

$$n \in \mathbb{N}$$

$$X = \mathbb{R}^n$$

$$Y = \{0, 1\}$$

$$L(r, \hat{y}) = -r\hat{y} + \log_p(1 + p^r)$$

$$f_\theta(x) : X \to \mathbb{R}, x \mapsto \langle \theta, x \rangle$$

$$S \subseteq X, S = \left\{x_1, x_2, ..., x_{|S|}\right\}, k = |S|$$

$$\inf_{(\theta,\varphi) \in X \times Y^S} \frac{\lambda}{2} \|\theta\|_2^2 + \sum_{x_i \in S} L\left(f_\theta(x_i), \varphi_{x_i}\right)$$

$$\theta := \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

Derivatives

$$\frac{\partial R + L\left(f_\theta(x_i), \varphi_{x_i}\right)}{\partial w_d} = \lambda w_d + x_{id}\left(-\varphi_{x_i} + \frac{1}{p^{-f_\theta(x_i)} + 1}\right)$$

$$\nabla_\theta R + L\left(f_\theta(x_i), \varphi_{x_i}\right) = \lambda\theta + \left(-\varphi_{x_i} + \frac{1}{p^{-f_\theta(x_i)} + 1}\right)x_i$$

$$\frac{\partial L\left(f_\theta(x_i), \varphi_{x_i}\right)}{\partial \varphi_{x_i}} = -f_\theta(x_i)$$

$$G = \nabla\left(R + \sum_{x_i \in S} L\left(f_\theta(x_i), \varphi_{x_i}\right)\right) = \begin{pmatrix} \lambda w_1 + \sum_{x_i \in S} x_{i1}\left(-\varphi_{x_i} + \frac{1}{p^{-f_\theta(x_i)}+1}\right) \\ \lambda w_2 + \sum_{x_i \in S} x_{i2}\left(-\varphi_{x_i} + \frac{1}{p^{-f_\theta(x_i)}+1}\right) \\ \vdots \\ \lambda w_n + \sum_{x_i \in S} x_{in}\left(-\varphi_{x_i} + \frac{1}{p^{-f_\theta(x_i)}+1}\right) \\ -f_\theta(x_1) \\ \vdots \\ -f_\theta(x_k) \end{pmatrix} = \begin{pmatrix} G_\theta \\ G_\varphi \end{pmatrix}$$