Konvexe Optimierung

Projekt

L_1 Loss

$$\min_{x\in\mathbb{R}^n} f_1(x) = \frac{1}{2}w^\intercal w + c \cdot \sum_{i=1}^{k} \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right] \tag{1}$$

$$\begin{aligned}
\frac{\partial}{\partial w} f_1(x) &= w + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w}\left[1 - y_i \cdot \left(w^\intercal x_i + b\right)\right] && (2)\\
&= w + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w}\left[-y_i \cdot \left(w^\intercal x_i + b\right)\right] && (3)\\
&= w + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w}\left[-y_i w^\intercal x_i + -y_i b\right] && (4)\\
&= w + c \cdot \sum_{i=1}^{k} -y_i x_i && (5)\\
& && (6)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial b} f_1(x) &= c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial b}\left[1 - y_i \cdot \left(w^\intercal x_i + b\right)\right] && (7)\\
&= c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial b}\left[-y_i w^\intercal x_i + -y_i b\right] && (8)\\
&= c \cdot \sum_{i=1}^{k} -y_i && (9)\\
& && (10)
\end{aligned}$$

L_2 Loss

$$\min_{x\in\mathbb{R}^n} f_2(x) = \frac{1}{2}w^\intercal w + c \cdot \sum_{i=1}^{k} \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}^2\right] \tag{11}$$

Logistic Loss

$$\min_{x\in\mathbb{R}^n} f_3(x) = \frac{1}{2}w^\intercal w + c \cdot \sum_{i=1}^{k} \left[\ln(1 + \exp(-y_i(w^\intercal x_i + b)))\right] \tag{12}$$

$$\begin{aligned}\frac{\partial}{\partial w} f_3(x) &= \frac{\partial}{\partial w}\left[\frac{1}{2} w^\intercal w\right] + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w}\left[\ln(1+\exp(-y_i(w^\intercal x_i + b)))\right] \\ &= w + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w}\left[\ln(1+\exp(-y_i(w^\intercal x_i + b)))\right]\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial x} f_3(x) &= \frac{\partial}{\partial x}\frac{1}{2} w^\intercal w + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial x}\left[\ln(1+\exp(-y_i(w^\intercal x_i + b)))\right] \\ &= c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial x}\left[\ln(1+\exp(-y_i\left(w^\intercal x_i + b\right)))\right] \\ &= c \cdot \sum_{i=1}^{k} \frac{\frac{\partial}{\partial x}\left[1+\exp(-y_i(w^\intercal x_i + b))\right]}{1+\exp(-y_i(w^\intercal x_i + b))} \\ &= c \cdot \sum_{i=1}^{k} \frac{\frac{\partial}{\partial x}\left[\exp(-y_i(w^\intercal x_i + b))\right]}{1+\exp(-y_i(w^\intercal x_i + b))} \\ &= c \cdot \sum_{i=1}^{k} \frac{\exp(-y_i(w^\intercal x_i + b))\frac{\partial}{\partial x}\left[-y_i(w^\intercal x_i + b)\right]}{1+\exp(-y_i(w^\intercal x_i + b))} \\ &= c \cdot \sum_{i=1}^{k} \frac{\exp(-y_i(w^\intercal x_i + b))\frac{\partial}{\partial x}\left[-y_i w^\intercal x_i - y_i b\right]}{1+\exp(-y_i(w^\intercal x_i + b))} \\ &= c \cdot \sum_{i=1}^{k} \frac{-y_i w \exp(-y_i(w^\intercal x_i + b))}{1+\exp(-y_i(w^\intercal x_i + b))}\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial x \partial x} f_3(x) &= \frac{\partial}{\partial x}\left[c \cdot \sum_{i=1}^{k} \frac{-y_i w \exp(-y_i(w^\intercal x_i + b))}{1+\exp(-y_i(w^\intercal x_i + b))}\right] \\ &= c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial x}\left[\frac{-y_i w \exp(-y_i(w^\intercal x_i + b))}{1+\exp(-y_i(w^\intercal x_i + b))}\right] = c \cdot \sum_{i=1}^{k} \frac{\frac{\partial}{\partial x} u(x_i) \cdot v(x_i) - u(x_i) \cdot \frac{\partial}{\partial x_i} v(x_i)}{v(x_i)^2}\end{aligned}$$

$$
\begin{aligned}
u(x) &= -y_i w \exp(-y_i(w^\intercal x_i + b)) \\
v(x) &= 1 + \exp(-y_i(w^\intercal x_i + b)) \\
\frac{\partial}{\partial x} u(x) &= \frac{\partial}{\partial x} \left[-y_i w \exp(-y_i(w^\intercal x_i + b)) \right] \\
&= -y_i w \exp(-y_i(w^\intercal x_i + b)) \frac{\partial}{\partial x} \left[-y_i(w^\intercal x_i + b) \right] \\
&= y_i^2 w^2 \exp(-y_i(w^\intercal x_i + b)) \\
\frac{\partial}{\partial x} v(x) &= \frac{\partial}{\partial x} \left[1 + \exp(-y_i(w^\intercal x_i + b)) \right] \\
&= \frac{\partial}{\partial x} \left[\exp(-y_i \left(w^\intercal x_i + b \right)) \right] \\
&= u(x)
\end{aligned}
$$

$$
\frac{\partial}{\partial x \partial x} f_3(x) = \frac{y_i^2 w^2 \exp(-y_i(w^\intercal x_i + b)) \cdot \left(1 + \exp(-y_i(w^\intercal x_i + b))\right) - \left(-y_i w \exp(-y_i(w^\intercal x_i + b))\right)^2}{(1 + \exp(-y_i(w^\intercal x_i + b)))^2}
$$