Konvexe Optimierung

Projekt

L_1 Loss

$$\min_{x\in\mathbb{R}^n} f_1(x) = \frac{1}{2}\left(w_1^2 + w_2^2\right) + c \cdot \sum_{i=1}^{k} \left[\max\left\{0, 1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right\}\right] \tag{1}$$

$$\begin{aligned}
\frac{\partial}{\partial w_1} f_1(x) &= \frac{\partial}{\partial w_1}\frac{1}{2}\left[w_1^2 + w_2^2\right] + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1}\left[1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right] \\
&= \frac{1}{2} 2 w_1 + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1}\left[-y_i w_1 x_1^{(i)} - y_i w_2 x_2^{(i)} - y_i b\right] \\
&= w_1 - c \cdot \sum_{i=1}^{k} y_i x_1^{(i)}
\end{aligned} \tag{2}$$

$$\frac{\partial}{\partial w_2} f_1(x) = w_2 - c \cdot \sum_{i=1}^{k} y_i x_2^{(i)} \tag{3}$$

$$\begin{aligned}
\frac{\partial}{\partial b} f_1(x) &= c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial b}\left[1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right] \\
&= c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial b}\left[-y_i w_1 x_1^{(i)} + w_2 x_2^{(i)} - y_i b\right] \\
&= -c \cdot \sum_{i=1}^{k} y_i
\end{aligned} \tag{4}$$

$$\mathcal{H}(f_1) = \begin{bmatrix} \frac{\partial^2}{\partial w_1 \partial w_1} f_1(x) & \frac{\partial^2}{\partial w_1 \partial w_2} f_1(x) & \frac{\partial^2}{\partial w_1 \partial b} f_1(x) \\ \frac{\partial^2}{\partial w_2 \partial w_1} f_1(x) & \frac{\partial^2}{\partial w_2 \partial w_2} f_1(x) & \frac{\partial^2}{\partial w_2 \partial b} f_1(x) \\ \frac{\partial^2}{\partial b \partial w_1} f_1(x) & \frac{\partial^2}{\partial b \partial w_2} f_1(x) & \frac{\partial^2}{\partial b \partial b} f_1(x) \end{bmatrix} \tag{5}$$

$$\frac{\partial^2}{\partial w_1 \partial w_1} f_1(x) = \frac{\partial}{\partial w_1} \left[w_1 - c \cdot \sum_{i=1}^{k} y_i x_1^{(i)} \right] = \frac{\partial}{\partial w_1} [w_1] = 1 \tag{6}$$

$$\frac{\partial^2}{\partial w_1 \partial w_2} f_1(x) = \frac{\partial}{\partial w_2} \left[w_1 - c \cdot \sum_{i=1}^{k} y_i x_1^{(i)} \right] = 0$$

$$\frac{\partial^2}{\partial w_1 \partial b} f_1(x) = \frac{\partial}{\partial b} \left[w_1 - c \cdot \sum_{i=1}^{k} y_i x_1^{(i)} \right] = 0$$

$$\frac{\partial^2}{\partial w_2 \partial w_1} f_1(x) = \frac{\partial}{\partial w_1} \left[w_2 - c \cdot \sum_{i=1}^{k} y_i x_2^{(i)} \right] = 0$$

$$\frac{\partial^2}{\partial w_2 \partial w_2} f_1(x) = \frac{\partial}{\partial w_2} \left[w_2 - c \cdot \sum_{i=1}^{k} y_i x_2^{(i)} \right] = \frac{\partial}{\partial w_2} [w_2] = 1 \tag{7}$$

$$\frac{\partial^2}{\partial w_2 \partial w_1 b} f_1(x) = \frac{\partial}{\partial b} \left[w_2 - c \cdot \sum_{i=1}^{k} y_i x_2^{(i)} \right] = 0$$

$$\frac{\partial^2}{\partial b \partial w_1} f_1(x) = \frac{\partial}{\partial w_1} \left[-c \cdot \sum_{i=1}^{k} y_i \right] = 0$$

$$\frac{\partial^2}{\partial b \partial w_2} f_1(x) = \frac{\partial}{\partial w_2} \left[-c \cdot \sum_{i=1}^{k} y_i \right] = 0$$

$$\frac{\partial^2}{\partial b \partial b} f_1(x) = \frac{\partial}{\partial b} \left[-c \cdot \sum_{i=1}^{k} y_i \right] = 0$$

$$\Rightarrow \mathcal{H}(f_2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{8}$$

L_2 Loss

$$\min_{x \in \mathbb{R}^n} f_2(x) = \frac{1}{2}\left(w_1^2 + w_2^2\right) + c \cdot \sum_{i=1}^{k} \left[\max\left\{0, 1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right\}^2 \right] \tag{9}$$

$$\begin{aligned}
\frac{\partial}{\partial w_1} f_2(x) &= \frac{1}{2} \frac{\partial}{\partial w_1} \left[w_1^2 + w_2^2\right] + c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1} \left[\left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right)^2 \right] \\
&= w_1 + c \cdot \sum_{i=1}^{k} 2\left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right) \frac{\partial}{\partial w_1} \left[1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right] \\
&= w_1 + c \cdot \sum_{i=1}^{k} 2\left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right) \frac{\partial}{\partial w_1} \left[-y_i w_1 x_1^{(i)} - y_i w_2 x_2^{(i)} - y_i b\right] \\
&= w_1 - c \cdot \sum_{i=1}^{k} 2 y_i x_1^{(i)} \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right)
\end{aligned} \tag{10}$$

$$\frac{\partial}{\partial w_2} f_2(x) = w_2 + c \cdot \sum_{i=1}^{k} 2 \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right) \frac{\partial}{\partial w_2} \left[-y_i w_1 x_1^{(i)} - y_i w_2 x_2^{(i)} - y_i b\right]$$

$$= w_2 - c \cdot \sum_{i=1}^{k} 2 y_i x_2^{(i)} \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right) \tag{11}$$

$$\tag{12}$$

$$\frac{\partial}{\partial b} f_2(x) = -c \cdot \sum_{i=1}^{k} 2 y_i \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right) \tag{13}$$

$$\tag{14}$$

$$\mathcal{H}(f_2) = \begin{bmatrix} \frac{\partial^2}{\partial w_1 \partial w_1} f_2(x) & \frac{\partial^2}{\partial w_1 \partial w_2} f_2(x) & \frac{\partial^2}{\partial w_1 \partial b} f_2(x) \\ \frac{\partial^2}{\partial w_2 \partial w_1} f_2(x) & \frac{\partial^2}{\partial w_2 \partial w_2} f_2(x) & \frac{\partial^2}{\partial w_2 \partial b} f_2(x) \\ \frac{\partial^2}{\partial b \partial w_1} f_2(x) & \frac{\partial^2}{\partial b \partial w_2} f_2(x) & \frac{\partial^2}{\partial b \partial b} f_2(x) \end{bmatrix} \tag{15}$$

$$\frac{\partial^2}{\partial w_1 \partial w_1} f_2(x) = \frac{\partial}{\partial w_1} \left[w_1 - c \cdot \sum_{i=1}^{k} 2 y_i x_1^{(i)} \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right)\right] \tag{16}$$

$$= 1 - c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1} \left[2 y_i x_1^{(i)} \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right)\right]$$

$$= 1 - c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1} \left[\left(2 y_i x_1^{(i)} - 2 y_i x_1^{(i)} y_i \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b\right)\right)\right]$$

$$= 1 - c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1} \left[-2 y_i x_1^{(i)} y_i w_1 x_1^{(i)}\right]$$

$$= 1 + c \cdot \sum_{i=1}^{k} 2 y_i x_1^{(i)} y_i x_1^{(i)}$$

$$\begin{aligned}
\frac{\partial^2}{\partial w_1 \partial w_2} f_2(x) &= -c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_2} \left[-2 y_i x_1^{(i)} y_i w_2 x_2^{(i)} \right] \\
&= c \cdot \sum_{i=1}^{k} 2 y_i x_1^{(i)} y_i x_2^{(i)}
\end{aligned} \tag{17}$$

$$\begin{aligned}
\frac{\partial^2}{\partial w_1 \partial b} f_2(x) &= -c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial b} \left[\left(2 y_i x_1^{(i)} - 2 y_i x_1^{(i)} y_i \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b \right) \right) \right] \\
&= c \cdot \sum_{i=1}^{k} 2 y_i x_1^{(i)} y_i
\end{aligned} \tag{18}$$

$$\tag{19}$$

$$\begin{aligned}
\frac{\partial^2}{\partial w_2 \partial w_1} f_2(x) &= \frac{\partial}{\partial w_1} \left[w_2 - c \cdot \sum_{i=1}^{k} 2 y_i x_2^{(i)} \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b \right) \right) \right] \\
&= -c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1} \left[2 y_i x_2^{(i)} \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b \right) \right) \right] \\
&= -c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1} \left[2 y_i x_2^{(i)} - 2 y_i x_2^{(i)} y_i w_1 x_1^{(i)} - 2 y_i x_2^{(i)} y_i w_2 x_2^{(i)} - 2 y_i x_2^{(i)} y_i b \right] \\
&= c \cdot \sum_{i=1}^{k} 2 y_i x_2^{(i)} y_i x_1^{(i)}
\end{aligned} \tag{20}$$

$$\tag{21}$$

$$\frac{\partial^2}{\partial w_2 \partial w_2} f_2(x) = 1 + c \cdot \sum_{i=1}^{k} 2 y_i x_2^{(i)} y_i x_2^{(i)} \tag{22}$$

$$\frac{\partial^2}{\partial w_2 \partial b} f_2(x) = c \cdot \sum_{i=1}^{k} 2 y_i x_2^{(i)} y_i \tag{23}$$

$$\begin{aligned}
\frac{\partial^2}{\partial b \partial w_1} f_2(x) &= \frac{\partial}{\partial w_1} \left[-c \cdot \sum_{i=1}^{k} 2 y_i \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b \right) \right) \right] \\
&= -c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_1} \left[2 y_i - 2 y_i y_i w_1 x_1^{(i)} - 2 y_i y_i w_2 x_2^{(i)} - 2 y_i y_i b \right] \\
&= c \cdot \sum_{i=1}^{k} 2 y_i y_i x_1^{(i)}
\end{aligned} \tag{24}$$

$$\begin{aligned}
\frac{\partial^2}{\partial b \partial w_2} f_2(x) &= \frac{\partial}{\partial w_2} \left[-c \cdot \sum_{i=1}^{k} 2 y_i \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b \right) \right) \right] \\
&= -c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_2} \left[2 y_i - 2 y_i y_i w_1 x_1^{(i)} - 2 y_i y_i w_2 x_2^{(i)} - 2 y_i y_i b \right] \\
&= c \cdot \sum_{i=1}^{k} 2 y_i y_i x_2^{(i)}
\end{aligned} \tag{25}$$

$$\begin{aligned}
\frac{\partial^2}{\partial b \partial b} f_2(x) &= \frac{\partial}{\partial w_2} \left[-c \cdot \sum_{i=1}^{k} 2 y_i \left(1 - y_i \cdot \left(w_1 x_1^{(i)} + w_2 x_2^{(i)} + b \right) \right) \right] \\
&= -c \cdot \sum_{i=1}^{k} \frac{\partial}{\partial w_2} \left[2 y_i - 2 y_i y_i w_1 x_1^{(i)} - 2 y_i y_i w_2 x_2^{(i)} - 2 y_i y_i b \right] \\
&= c \cdot \sum_{i=1}^{k} 2 y_i y_i
\end{aligned} \tag{26}$$

$$\Rightarrow \mathcal{H}(f_2) = \begin{bmatrix} 1 + c \cdot \sum_{i=1}^{k} 2y_i x_1^{(i)} y_i x_1^{(i)} & c \cdot \sum_{i=1}^{k} 2y_i x_1^{(i)} y_i x_2^{(i)} & c \cdot \sum_{i=1}^{k} 2y_i x_1^{(i)} y_i \\ c \cdot \sum_{i=1}^{k} 2y_i x_2^{(i)} y_i x_1^{(i)} & 1 + c \cdot \sum_{i=1}^{k} 2y_i x_2^{(i)} y_i x_2^{(i)} & c \cdot \sum_{i=1}^{k} 2y_i x_2^{(i)} y_i \\ c \cdot \sum_{i=1}^{k} 2y_i y_i x_1^{(i)} & c \cdot \sum_{i=1}^{k} 2y_i y_i x_2^{(i)} & c \cdot \sum_{i=1}^{k} 2y_i y_i \end{bmatrix} \tag{27}$$

Logistic Loss

$$\min_{x \in \mathbb{R}^n} f_3(x) = \frac{1}{2} w^\mathsf{T} w + c \cdot \sum_{i=1}^{k} \left[\ln(1 + \exp(-y_i(w^\mathsf{T} x_i + b))) \right] \tag{28}$$

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

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

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

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

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