Konvexe Optimierung

Projekt

L_1 Loss

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

Dual Form

$$\begin{aligned} \max_{\alpha} \min_{w,b} f_1^{\alpha}(x) &= \frac{1}{2} w^\intercal w + c \cdot \sum_{i=1}^{k} \alpha_i \left[\max \left\{ 0, 1 - y_i \cdot \left(w^\intercal x_i + b \right) \right\} \right] \\ &= \frac{1}{2} w^\intercal w + c \cdot \sum_{i=1}^{k} \alpha_i \left(1 - y_i \cdot \left(w^\intercal x_i + b \right) \right) && (2) \\ & && (3) \end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial w} f_3^\alpha(x) &= \frac{\partial}{\partial w}\left[\frac{1}{2} w^\intercal w + c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right]\right] \\
&= \frac{\partial}{\partial w}\left[\frac{1}{2} w^\intercal w\right] + \frac{\partial}{\partial w}\left[c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right]\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[1 - y_i \cdot \left(w^\intercal x_i + b\right)\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[1 - y_i \cdot \left(w^\intercal x_i + b\right)\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[-y_i \cdot \left(w^\intercal x_i + b\right)\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[\left(-y_i w^\intercal x_i + -y_i b\right)\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[-y_i w^\intercal x_i + -y_i b\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i - y_i x_i \\
&\Rightarrow w = c \cdot \sum_{i=1}^{k} \alpha_i - y_i x_i
\end{aligned} \tag{4}
$$

$$
\begin{aligned}
\frac{\partial}{\partial b} f_3^\alpha(x) &= \frac{\partial}{\partial b}\left[\frac{1}{2} w^\intercal w + c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right]\right] \\
&= \frac{\partial}{\partial b}\left[c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right]\right] \\
&= c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial b} \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right] \\
&= c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial b} \left[1 - y_i \cdot \left(w^\intercal x_i + b\right)\right] \\
&= c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial b} \left[-y_i \cdot \left(w^\intercal x_i + b\right)\right] \\
&= c \cdot \sum_{i=1}^{k} \alpha_i - y_i
\end{aligned}
$$

$$
\begin{aligned}
\max_{\alpha} \min_{w,b} f_1^\alpha(x) &= \frac{1}{2} w^\intercal w + c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\intercal x_i + b\right)\right\}\right] \\
&= \frac{1}{2} w^\intercal w + c \cdot \sum_{i=1}^{k} \alpha_i \left(1 - y_i \cdot \left(w^\intercal x_i + b\right)\right) && (5) \\
&= \frac{1}{2}\left(c \cdot \sum_{i=1}^{k} \alpha_i - y_i x_i\right)^\intercal \left(c \cdot \sum_{i=1}^{k} \alpha_i - y_i x_i\right) + c \cdot \sum_{i=1}^{k} \alpha_i \left(1 - y_i \cdot \left(\left(c \cdot \sum_{i=1}^{k} \alpha_i - y_i x_i\right)^\intercal x_i + b\right)\right) && (6) \\
& && (7)
\end{aligned}
$$

L_2 Loss

$$\min_{w,b} f_2(x) = \frac{1}{2}w^\top w + c \cdot \sum_{i=1}^{k} \left[\max\left\{0, 1 - y_i \cdot \left(w^\top x_i + b\right)\right\}^2\right] \tag{8}$$

Dual Form

$$\max_{\alpha} \min_{w,b} f_2^\alpha(x) = \frac{1}{2}w^\top w + c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\top x_i + b\right)\right\}^2\right]$$

$$\begin{aligned}
\frac{\partial}{\partial w} f_2^\alpha(x) &= \frac{\partial}{\partial w}\left[\frac{1}{2}w^\top w + c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\top x_i + b\right)\right\}^2\right]\right] \\
&= \frac{\partial}{\partial w}\left[\frac{1}{2}w^\top w\right] + \frac{\partial}{\partial w}\left[c \cdot \sum_{i=1}^{k} \alpha_i \left[\max\left\{0, 1 - y_i \cdot \left(w^\top x_i + b\right)\right\}^2\right]\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[\max\left\{0, 1 - y_i \cdot \left(w^\top x_i + b\right)\right\}^2\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[\left(1 - y_i \cdot \left(w^\top x_i + b\right)\right)^2\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[1 - 2\left(y_i \cdot \left(w^\top x_i + b\right)\right) + y_i^2 \cdot \left(w^\top x_i + b\right)^2\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i \frac{\partial}{\partial w}\left[1 - 2y_i w^\top x_i - 2y_i b + y_i^2 w^\top w x_i^2 + y_i^2 2 w^\top x_i b + y_i^2 b^2\right] \\
&= w + c \cdot \sum_{i=1}^{k} \alpha_i 2 y_i x_i + 2 y_i^2 w^\top x_i^2 + y_i^2 2 w^\top x_i b
\end{aligned} \tag{9}$$

Logistic Loss

$$\min_{w,b} f_3(x) = \frac{1}{2}w^\intercal w + c \cdot \sum_{i=1}^{k} \left[\ln(1 + \exp(-y_i(w^\intercal x_i + b)))\right] \tag{10}$$

Dual Form

$$\max_{\alpha} \min_{w,b} f_3^\alpha(x) = \frac{1}{2}w^\intercal w + c \cdot \sum_{i=1}^{k} \alpha_i \left[\ln(1 + \exp(-y_i(w^\intercal x_i + b)))\right]$$

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