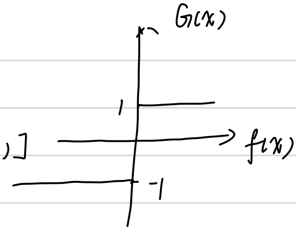
Adaboost

模型：加法模型

最终分类器：$G(x) = \mathrm{sign}[f(x)]$



损失函数：指数损失函数

二分类问题，使用指数损失函数

$L(y, f(x)) = e^{-yf(x)}$

G(x)分类正确，f(x)与y同号 $L(y, f(x)) \leq 1$

G(x)分类错误，f(x)与y异号 $L(y, f(x)) > 1$

分类正确 — loss小

分类错误 — loss大

将loss当作训练数据的权值

$\overline{w}_{mi} = e^{-y_i f_{m-1}(x_i)}$

单个样本损失函数：

$$L(y, f_m(x)) = e^{-y f_m(x)} = e^{-y \sum_{m=1}^{M} \alpha_m G_m(x)} = e^{-y[f_{m-1}(x) + \alpha_m G_m(x)]}$$

总体损失函数（包含所有样本）$\sum_{i=1}^{N} e^{[-y_i f_{m-1}(x_i) + \alpha G(x_i)]}$

[优化方法]：前向分步算法 第m轮

$$(\alpha_m, G_m(x)) = \arg\min_{\alpha, G} \sum_{i=1}^{N} e^{[-y_i f_{m-1}(x_i) + \alpha G(x_i)]}$$

$$= \arg\min \sum_{i=1}^{N} e^{-y_i f_{m-1}(x_i)} \cdot e^{-y_i \alpha G(x_i)}$$

$$= \arg\min \sum_{i=1}^{N} \overline{w}_{mi} \cdot e^{[-y_i \alpha G(x_i)]}$$

2种取值可能

$\begin{cases} G(x_i) = y_i \\ G(x_i) \neq y_i \end{cases}$

同号　异号

$$= \arg\min \Big(\sum_{y_i = G(x_i)} \overline{w}_{mi} \cdot e^{-y_i \alpha G(x_i)} + \sum_{y_i \neq G(x_i)} \overline{w}_{mi} \cdot e^{-y_i \alpha G(x_i)} \Big)$$

$$= \arg\min \Big(\sum_{y_i = G(x_i)} \overline{w}_{mi} \cdot e^{-\alpha} + \sum_{y_i \neq G(x_i)} \overline{w}_{mi} e^{\alpha} \Big)$$

$$= \arg\min_{\alpha, a} \Big(e^{-\alpha} \sum_{y_i = G(x_i)} \overline{w}_{mi} + e^{\alpha} \sum_{y_i \neq G(x_i)} \overline{w}_{mi} \Big)$$ 其中，$\overline{w}_{mi} = e^{-y_i f_{m-1}(x_i)}$

第m轮求解:

1.优化 $G_m(x)$ = 使得 $G_m(x)$ 的分类误差率最小

$$\alpha_m^*(x) = \arg\min \sum_{i=1}^{N} \bar{w}_{mi} I(y_i \neq G(x_i))$$

损失函数作为样本权值

↳ 当 $y_i \neq G(x_i)$ 成立时 $I(y_i \neq G(x_i)) = 1$

2.优化 α_m

分类器 $G_m(x)$ 的权重系数

$$原式 = \arg\min_{\alpha_m} \left(e^{-\alpha_m} \sum_{y_i = G(x_i)} \bar{w}_{mi} + e^{\alpha_m} \sum_{y_i \neq G(x_i)} \bar{w}_{mi}\right)$$

$$\Rightarrow \arg\min_{\alpha_m} \left(e^{-\alpha_m} \sum_{y_i = G(x_i)} \bar{w}_{mi} + e^{-\alpha_m} \sum_{y_i \neq G(x_i)} \bar{w}_{mi} + e^{\alpha_m} \sum_{y_i \neq G(x_i)} \bar{w}_{mi} - e^{-\alpha_m} \sum_{y_i \neq G(x_i)} \bar{w}_{mi}\right)$$

$$\Rightarrow \arg\min \left(e^{-\alpha_m} \sum_{i=1}^{N} \bar{w}_{mi} + (e^{\alpha_m} - e^{-\alpha_m}) \sum_{y_i \neq G(x_i)} \bar{w}_{mi}\right) \quad 式(1)$$

已优化步骤:

(1) 式(1)为 α_m 求导

注: 式(1)中, 只有 α_m 为变量, 其余均为定值

$$\frac{\partial \left(e^{-\alpha_m} \sum_{i=1}^{N} \bar{w}_{mi} + (e^{\alpha_m} - e^{-\alpha_m}) \sum_{y_i \neq G(x_i)} \bar{w}_{mi}\right)}{\partial \alpha_m}$$

$$= -e^{-\alpha_m} \sum_{i=1}^{N} \bar{w}_{mi} + e^{\alpha_m} \sum_{y_i \neq G(x_i)} \bar{w}_{mi} + e^{-\alpha_m} \sum_{y_i \neq G(x_i)} \bar{w}_{mi}$$

$$= -e^{-\alpha_m} \sum_{y_i = G_i(x_i)} \bar{w}_{mi} + e^{\alpha_m} \sum_{y_i \neq G(x_i)} \bar{w}_{mi} \quad 式(2)$$

② 导数为 0

令式(2)=0

$$-e^{-\alpha_m}\sum_{y_i=G(x_i)}\overline{w}_{mi} + e^{\alpha_m}\sum_{y_i\neq G(x_i)}\overline{w}_{mi}=0$$

$$\Rightarrow e^{\alpha_m}\sum_{y_i\neq G(x_i)}\overline{w}_{mi} = e^{-\alpha_m}\sum_{y_i=G(x_i)}\overline{w}_{mi}$$

2边同时取对数

$$\ln\Big(e^{\alpha_m}\sum_{y_i\neq G(x_i)}\overline{w}_{mi}\Big) = \ln\Big(e^{-\alpha_m}\sum_{y_i=G(x_i)}\overline{w}_{mi}\Big)$$

$$\ln e^{\alpha_m}+\ln\Big(\sum_{y_i\neq G(x_i)}\overline{w}_{mi}\Big) = \ln e^{-\alpha_m}+\ln\Big(\sum_{y_i=G(x_i)}\overline{w}_{mi}\Big)$$

$$2\alpha_m = \ln\Big(\sum_{y_i=G(x_i)}\overline{w}_{mi}\Big)-\ln\Big(\sum_{y_i\neq G(x_i)}\overline{w}_{mi}\Big)$$

$$2\alpha_m = \ln\frac{\sum_{y_i=G(x_i)}\overline{w}_{mi}}{\sum_{y_i\neq G(x_i)}\overline{w}_{mi}}$$ (分类正确 = all − 分类错误)

$$\alpha_m=\frac{1}{2}\ln\left(\frac{\sum_{i=1}^{N}\overline{w}_{mi}-\sum_{y_i\neq G(x_i)}\overline{w}_{mi}}{\sum_{y_i\neq G(x_i)}\overline{w}_{mi}}\right)$$

③ 转换

$$\alpha_m=\frac{1}{2}\ln\left(\frac{\sum_{i=1}^{N}\overline{w}_{mi}-\sum_{y_i\neq G(x_i)}\overline{w}_{mi}}{\sum_{y_i\neq G(x_i)}\overline{w}_{mi}}\right)$$

上下同除所有样本权重之和

$$=\frac{1}{2}\ln\left(\frac{\frac{\sum_{i=1}^{N}\overline{w}_{mi}}{\sum_{i=1}^{N}\overline{w}_{mi}}-\frac{\sum_{y_i\neq G(x_i)}\overline{w}_{mi}}{\sum_{i=1}^{N}\overline{w}_{mi}}}{\frac{\sum_{y_i\neq G(x_i)}\overline{w}_{mi}}{\sum_{i=1}^{N}\overline{w}_{mi}}}\right)$$

$$=\frac{1}{2}\ln\left(\frac{1-e_m}{e_m}\right)$$

分类误差率 $\frac{\text{error}}{\text{all}}$

$\therefore$ 最终得到 $\alpha_m = \frac{1}{2}\log\frac{1-e_m}{e_m}$

(e_m为分类误差率，e_m越小，α_m越大)

$$f(x) = \sum_{m=1}^{M} \alpha_m G_m(x) = \alpha_1 G_1(x) + \cdots + \alpha_m G_m(x)$$

由损失函数确定训练集权重

由 $G_1(x)$ 的"分类误差率"决定

3. 前向更新 $f_m(x)$ $\quad f_m(x) = f_{m-1}(x) + \alpha_m G_m(x)$

4. 更新训练数据权值 $\quad \overline{w}_{m+1,i}$ (m+1轮) $\quad \overline{w}_{mi} = e^{-y_i f_{m-1}(x_i)}$

(未做归一化) $\quad \overline{w}_{mi} \triangleq \begin{cases} \frac{1}{N} & m=1 \\ \overline{w}_{m-1,i} \cdot e^{-y_i \alpha_{m-1} G_{m-1}(x_i)} & m>1 \end{cases}$

推导：由 $\overline{w}_{mi} = e^{-y_i f_{m-1}(x_i)}$

$$\begin{aligned}
\overline{w}_{m+1,i} &= e^{-y_i f_m(x_i)} \\
&= e^{-y_i [f_{m-1}(x_i) + \alpha_m G_m(x_i)]} \\
&= e^{-y_i f_{m-1}(x_i)} \cdot e^{-y_i \alpha_m G_m(x_i)} \\
&= \overline{w}_{m,i} \cdot e^{-y_i \alpha_m G_m(x_i)}
\end{aligned}$$