$$y_i \in \{0,1\}, p_i = \frac{e^{pred_i}}{1+e^{pred_i}} \tag{1}$$

$$\frac{\partial p}{\partial pred} = p(1-p) \tag{2}$$

1 Initial

$$\mathrm{L}(Y, pred) = -\sum(y_i * pred - log(1+e^{pred})) \tag{3}$$

$$g = \frac{\partial L}{\partial pred} = -\sum(y_i - p) \tag{4}$$

$$\arg\min_{pred} \mathrm{L}(Y, pred) \Rightarrow \frac{\partial L}{\partial pred} = 0 \Rightarrow pred = log\frac{1_(yi=1)}{1_(yi=0)} = log\frac{count(yi=1)}{count(yi=0)} \tag{5}$$

2 Build Tree, Focus on the leaf , a.k.a terminal region

NOTE predi is the begin predict state of every sample on the tree beginning.delta is the leaf value.REMEMBER: We have predi , we need to get delta.

$$Y = \{y_i\}, \arg\min_{\delta} \mathrm{L}(Y, \{pred_i + \delta\}) \tag{6}$$

$$\Rightarrow using\ the\ second\ taylor\ expansion\ at\ pred_i,\ then\ \frac{\partial L}{\partial \delta} = 0 \tag{7}$$

$$\Rightarrow \delta = -g/H, g_i = \frac{\partial L}{\partial pred_i} = y_i - p_i, H = \frac{\partial g_i}{\partial pred_i} = -\frac{\partial p_i}{\partial pred_i} = -p_i(1-p_i) \tag{8}$$

$$\Rightarrow final, \delta = \frac{y_i - p_i}{p_i(1-p_i)} \tag{9}$$

On predicting, use the followiong :

$$p_i = \frac{e^{pred_i}}{1+e^{pred_i}}, \tag{10}$$

$$where\ pred_i = pred_{t0} + \lambda(pred_{t1} + ... + pred_{tn});\ ti\ is\ tree_i, \lambda\ is\ learning-rate. \tag{11}$$

3 Inspire

- Pick any leaf in the first tree, it maybe contains 10 or 100 train-samples, How many y-values does the leaf contains? How about the second tree?
- When we have 100 trees ,every tree have 8 leaf , How many values we can predict in permutation? (Predict-Sapce)