Algorithm

Monday, September 20, 2021 8:23 AM

$\rightarrow \{(x_i, y_i)\}_{i=1}^{n}$ $n=3$ x_i y_i

Input: training set $\{(x_i, y_i)\}_{i=1}^{n}$, a differentiable loss function $L(y, F(x))$, number of iterations M.

1. Initialize $f_0(x) = \arg\min_\gamma \sum_{i=1}^{N} L(y_i, \gamma)$.
2. For $m = 1$ to $\boxed{M}$:
   (a) For $i = 1, 2, \ldots, N$ compute

   residual / pseudo-residual

   row no

   $i / m \rightarrow$

$$r_{im} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f=f_{m-1}}.$$

   (b) Fit a regression tree to the targets r_{im} giving terminal regions $R_{jm},\ j = 1, 2, \ldots, J_m$.

   (c) For $j = 1, 2, \ldots, J_m$ compute

$$\boxed{\gamma_{jm}} = \arg\min_{\gamma} \sum_{x_i \in R_{jm}} L\left(y_i, f_{m-1}(x_i) + \gamma\right).$$

   (d) Update $f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm})$.
3. Output $\hat{f}(x) = \boxed{f_M(x)}$.

$\boxed{1}$ $f_1(x) = f_0(x) + dT$

$f_3(x) = f_2(x) + dT3$

$f_1(x) + dT2$

$f_0(x) + dT1$

$f_2(x) = f_1(x) + dT2$

$f_0(x) + dT$

$\rightarrow$ recursion

$f_4(x) = f_0(x) + \ldots$

$f_1(x)$ $f_2(x)$

Additive Modelling

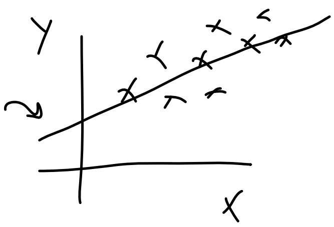
Monday, September 20, 2021 8:28 AM

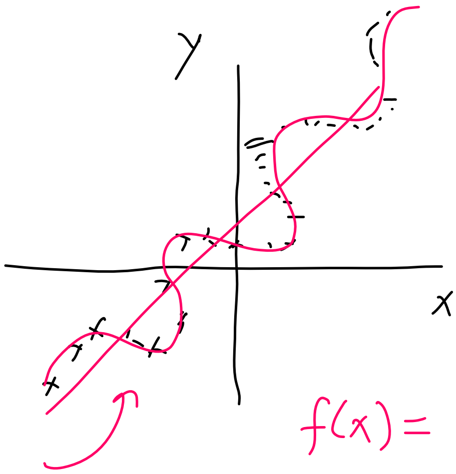
$X \mid y \longrightarrow f(\ ) \Rightarrow y = f(x)$

$x\ y$

$y = f(x_1, x_2, x_3)$

$x_1\ x_2\ x_3 \mid y$





$f(x) = x + \sin x$

additive

$F(x) = f_0(x) + f_1(x) + f_2(x) + \ldots$

DT PT

$y = x \quad y = \sin(x)$

Explanation

Monday, September 20, 2021 8:25 AM

$$L = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

actual pred

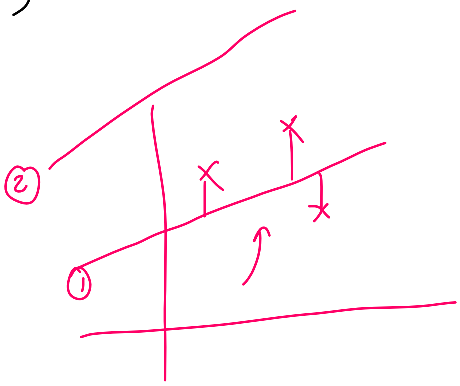
ls

$$L = \frac{1}{2} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \rightarrow \text{diff}$$

$\frac{1}{2}$

$$L(y, F(x)) \rightarrow L(y, \hat{y})$$

$\hat{y}$



① = 10 = 5 ✓

② = 20 = 10

$$y = f(x) \sim$$

$$f(x) = f_0(x) + f_1(x) + f_2(x) + \dots + f_n(x)$$

DT

$$L = \frac{1}{2} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$f_0(x) = \operatorname{argmin}_{\gamma} \sum_{i=1}^{n} L(y_i, \gamma)$$

$$f_0(x) = \operatorname{argmin}_{\gamma} \boxed{\frac{1}{2} \sum_{i=1}^{n} (y_i - \gamma)^2}$$

γ

$$\frac{d f_0(x)}{d\gamma} = \frac{d}{d\gamma} \frac{1}{2} \sum_{i=1}^{n} (y_i - \gamma)^2 = \frac{1}{2} \sum_{i=1}^{n} \frac{d}{d\gamma} (y_i - \gamma)^2$$

$$\sum_{i=1}^{n} (y_i - \gamma) \frac{d}{d\gamma} (y_i - \gamma) = -\sum_{i=1}^{n} (y_i - \gamma) = 0$$

$$\sum_{i=1}^{n} (\gamma - y_i) = 0$$

$$\sum_{i=1}^{3} (\gamma - y_i) = 0 \Rightarrow (\gamma - 192) + (\gamma - 144) + (\gamma - 91) = 0$$

$$\sum_{i=1} (\gamma - \underline{y_i}) = 0 \Rightarrow (\gamma - 192) + (\gamma - 144) + (\gamma \ldots$$

$$3\gamma = 192 + 144 + 91$$

mean

$F_m(x) = f_0(x)$ → mean of output

$$\boxed{\gamma = \frac{192 + 144 + 91}{3}}$$

$$F(x) = \underline{f_0(x)} + \underline{f_1(x)} + \underline{f_2(x)} + \cdots + \underline{f_m(x)}$$

DT

mean (leaf)

$\boxed{m = 1}$

$$\underline{r_{im}} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f = f_{m-1}}$$

$$r_{i1} = -\left[\frac{\partial \boxed{L(y_i, f(x_i))}}{\underline{\partial f(x_i)}}\right]_{f = f_0}$$

3 row

r_{11} r_{21} r_{31}

$\underline{\hat{y}_i = f(x_i)}$

$$r_{i1} = -\left[\frac{\partial}{\partial \hat{y}_i} L(y_i, \hat{y}_i)\right]_{f = f_0}$$

$$L = \frac{1}{2}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$r_{i1} = -\left[\frac{\partial}{\partial \hat{y}_i} \frac{1}{2} \underline{(y_i - \hat{y}_i)}^2\right]_{f = f_0}$$

-1

$$= \left[(y_i - \hat{y}_i)\right]_{f = f_0} = \left[(y_i - f(x_i))\right]_{f = \underline{f_0}}$$

$$= \left[(y_i - y_i) \right]_{f=f_0} = \left[(y_i - f^{(m-1)}) \right]_{f=f_0}$$

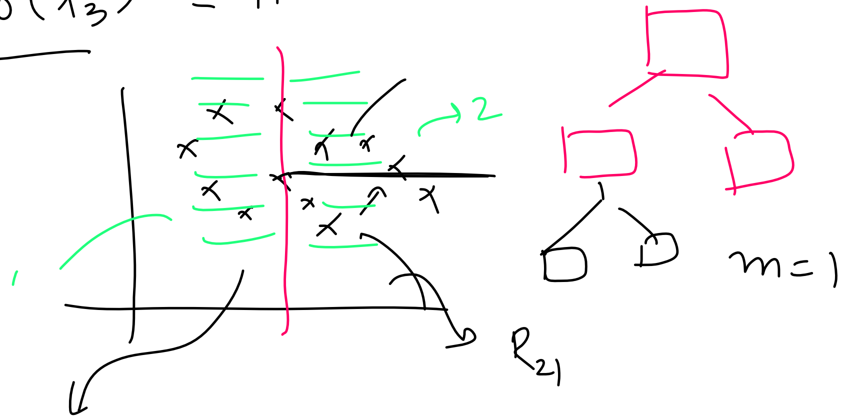
$$r_{i1} = (y_i - f_0(x_i))$$

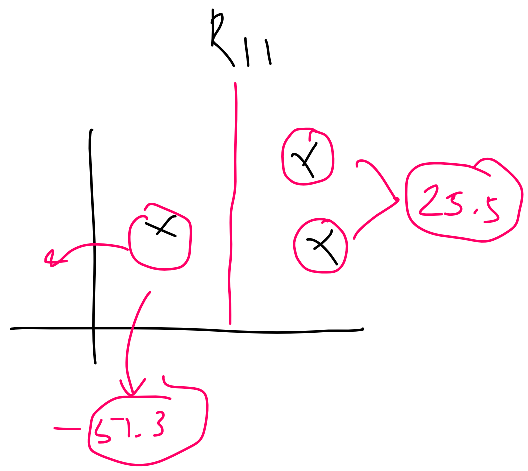
$$r_{11} = y_1 - f_0(x_1) = 192 - 142 =$$

$$r_{21} = y_2 - f_0(x_2) = 144 - 142 =$$

$$r_{31} = y_3 - f_0(x_3) = 91 - 142 =$$

x | y





$$\gamma_{jm} = \underset{\gamma}{\operatorname{argmin}} \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma)$$

$$\gamma_{j1} \Rightarrow \gamma_{11} = \underset{\gamma}{\operatorname{argmin}} \sum_{x_i \in R_{11}} L(y_i, f_{m-1}(x_i) + \gamma)$$

γ_{11} γ_{21}

$$\gamma_{11} \qquad \gamma_{21}$$

$$\gamma_{11} = \underset{\gamma}{\operatorname{argmin}} \frac{1}{2}\left(y_i - (f_0(x_i) + \gamma\right)^2$$

$$\frac{dL}{d\gamma} = \frac{1}{\cancel{2}} \times \cancel{2}\left(y_i - (f_0(x) + \gamma)\right) \frac{d}{d\gamma}\left(y_i - f_0(x) - \gamma\right) = 0$$

$$= -(y_i - f_0(x) - \gamma) = 0$$

$$= y_i - f_0(x) - \gamma = 0$$

$$\gamma_{11} = 91 - 142 - \gamma = 0$$

$$\boxed{\gamma = 91 - 142 = -51}$$

$$\gamma_{21} = \underset{\gamma}{\operatorname{argmin}} \sum_{x_i \in R_{21}} L(y_i, f_0(x_i) + \gamma)$$

$$= \underset{\gamma}{\operatorname{argmin}} \frac{1}{2} \sum_{i=1}^{2} \left(y_i - (f_0(x_i) + \gamma\right)^2$$

$$= -\sum_{i=1}^{2} (y_i - f_0(x_i) - \gamma) = 0$$

$$= \sum_{i=1}^{2} (y_i - f_0(x_i) - \gamma) = 0$$

$$= y_1 - f_0(x_1) - \gamma + y_2 - f_0(x_2) - \gamma = 0$$

336
284

$$= 192 - 142 - \gamma + 144 - 142 - \gamma = 0$$

$$\underbrace{336 - 284}$$

$$52 - 2\gamma = 0$$

$$\gamma = \cancel{\frac{52}{2}} \quad 26$$