# Probabilistic Boosting Tree

| | | | | |
|---|---|---|---|---|
| LSS GamLSS/GamboostLSS [1] LSS location/shape/skewness | | R | gamlss | |
| Gaussian Process Regression[2] / Distributional learning [3] | | | | |
| Quantile Regression | | | | quantile |

Probabilistic Boosting Tree

$y$ $x$ $y$ $p(y|x)$ boosting tree $T_1$ $l_i$ $p_{l_i}(y)$

$$p_{T_1}(y|x)$$

d Boosting

$$q(y|x) = \frac{\prod_{i=0}^{d-1} p_{T_i}(y|x)}{C}$$

$C$

$$q_{T_m}(y|x) = \frac{p_{T_m}(y|x) \prod_{i=0}^{m-1} p_{T_i}(y|x)}{C}$$

$q_{T_{m-1}}(y|x) = \prod_{i=0}^{m-1} p_{T_i}(y|x)$ $p_{T_m}(y|x)$ $q_{T_m}$
Gamma Gamma 0

$$q_{T_m}(y|x) = \frac{1}{\Gamma(k_m){\theta_m}^{k_m}} y^{k_m - 1} e^{-\frac{y}{\theta_m}}$$

$k_{m-1} \to k_m$ $\theta_{m-1} \to \theta_m$

$k_m = k_{m-1} + \Delta k$ $\theta_m = \theta_{m-1} + \Delta\theta$ $\Delta k$ $\Delta\theta$

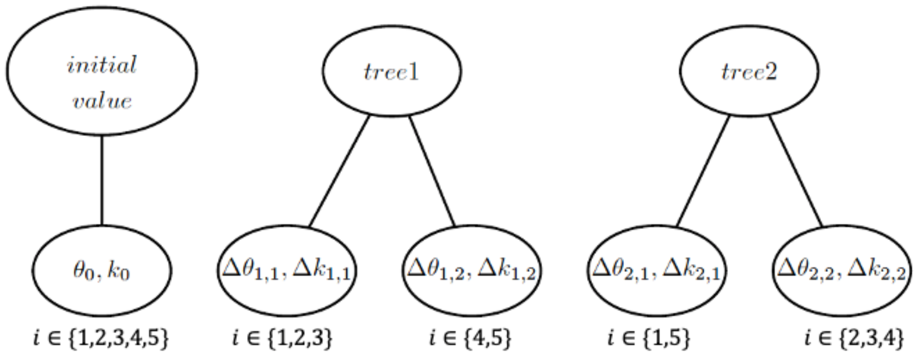


Figure 1: PBTree

- $\theta_{2,1} = \theta_0 + \Delta\theta_{1,1} + \Delta\theta_{2,1}$
- $\theta_{2,2} = \theta_0 + \Delta\theta_{1,1} + \Delta\theta_{2,2}$
- $\theta_{2,5} = \theta_0 + \Delta\theta_{1,2} + \Delta\theta_{2,1}$

n $m$

$$L_m = -log(\prod_{i=1}^{n} p(y_i|\theta_{m-1,i} + \Delta\theta_l, k_{m-1,i} + \Delta k_l)), \quad i \in S_{m,l}$$

$$\theta_{m-1,i} = \theta_0 + \sum_{j=0}^{m-1} \theta_l$$

$$k_{m-1,i} = k_0 + \sum_{j=0}^{m-1} k_l$$

$i \in S_{m,l}$ $i$ $m$ $l$

$\Delta\theta$ $\Delta k$

$$\arg\min_{S,\Theta,K} L_m$$

$$L_m = -log(\prod_{i=0}^{n} p(y_i|\theta_{m-1,i} + \Delta\theta_{l,\{i\in S_{m,l}\}}, k_{m-1,i} + \Delta k_{l,\{i\in S_{m,l}\}})) = -\sum_{i=0}^{n} log(p(y_i|\theta_{m-1,i} + \Delta\theta_{l,\{i\in S_{m,l}\}}, k_{m-1,i}$$

$m$ $l$ $(x_i, y_i), i = \{1...s\}$

$$\frac{\partial L_m}{\partial k_{m,l}} = -\sum_{i=1}^{s} (log(\frac{y_i}{\theta_{m-1,l}}) + F(k_{m-1,l}))$$

$$\frac{\partial L_m}{\partial \theta_{m,l}} = -\sum_{i=1}^{s} (\frac{y_i}{\theta^2_{m-1,l}} - \frac{k_{m-1,l}}{\theta_{m-1,l}})$$

$\eta_1, \eta_2$

$$\Delta k_{m,l} = \eta_1 \frac{\partial L_m}{\partial k_{m,l}}$$

$$\Delta \theta_{m,l} = \eta_2 \frac{\partial L_m}{\partial \theta_{m,l}}$$