

Drawing for PFV Interview 2.

Grid Search \rightarrow Bayesian Optimization - based on Gaussian process models (low-dimension)
 \searrow Random Forest (high-dimension)

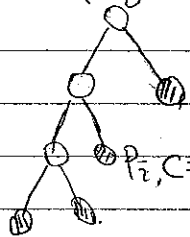
$\hat{f} \leftarrow$ predictive model of f

$X(\theta_U) = \{\theta' \in \Theta \mid \theta'_U = \theta_U\}$ Extension Set.

$$\|S\| = \begin{cases} 1 & (S = \phi) \\ |S| & (|S| < \infty) \\ r-1 & (S = [2, r]) \\ \prod_{i=1}^k \|S_i\| & (S = S_1 \times \dots \times S_k) \end{cases}$$

Random Forests は regression trees の集まり

再帰木 (regression tree)



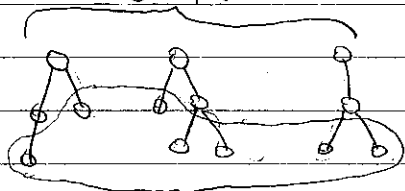
各葉に, $P_i = \theta_1^{(i)} \times \dots \times \theta_n^{(i)}$ に対して, Θ を分割している。

$$\hat{f}(\theta) = \sum_{P_i \in \mathcal{P}} \mathbb{I}(\theta \in P_i) \cdot C_i$$

$$\mathbb{I}(\theta \in P_i) = \begin{cases} 1 & \theta \in P_i \\ 0 & \theta \notin P_i \end{cases}$$

Random Forest においては $\hat{f}(\theta)$ を計算するとき, 各 regression tree において $\hat{f}(\theta)$ を計算し, その平均を取る。

B trees



L leaves

Preprocessing Phase \rightarrow Prediction Phase
 各木において P_i の構築 各 $\hat{f}_i(\theta_U)$ の計算