Branch and Bound

-Statistical Learning and Data Mining-

Lecturer: Darren Homrighausen, PhD

Branch and bound

Let $M = \{M_1, \ldots, M_K\}$ be the set of all possible solutions and a partition comprised of branches, respectively.

(Statistically, we think of M as the set of all possible models and for all subsets, M_1 is the set of all size 1 models, M_2 is the set of all size 2 models, ...)

Let F be the objective function we wish to maximize and:

$$m_* = \max_{m \in M} F(m)$$

For each M_k, define

$$m_k = \max_{m \in M_k} F(m)$$

and let $\underline{m}_k, \overline{m}_k$ be a bracket such that

$$\underline{m}_k \leq m_k \leq \overline{m}_k$$

(Note that m_k is in general not explicitly constructed)

Then

$$\max_k \underline{m}_k = \underline{m} \leq m_* \leq \overline{m} = \max_k \overline{m}_k$$

Branch and bound

The main realization is that the branch M_k does not need to be explored if either of the following occur

I. Bound

$$\overline{m}_k \leq \underline{m}$$

II. Optimality

$$\max_{m \in M_k} F(m) \text{ has been found}$$

The two main questions remain:

1. How to choose the partition(s)?
2. How to form the upper/lower bounds?

These are very case specific. Let's return to model selection

Branch and bound for model selection

Let's suppose we set[1]

$$F(m) = \hat{R}(\hat{\beta}_m) + 2|m|$$

For the M_k, let

$m_{k,inf}$ be the largest model contained[2] in every model in M_k

$m_{k,sup}$ be a smallest model that contains every model in M_k

[1] Note: we are trying to minimize F, not maximize

[2] This does not have to be in M_k

Branch and bound for model selection

Example: Let $x_1, \ldots, x_5$ be covariates

$$M = \cup_{k=1}^{3} M_k,$$

where

$$\begin{aligned} M_1 &= \{\{x_1, x_3\}, \{x_2\}\}, \\ M_2 &= \{\{x_2, x_3, x_4\}, \{x_3, x_4\}\}, \\ M_3 &= \{\{x_3, x_5\}, \{x_3\}\}, \end{aligned}$$

Branch and bound for model selection

Example: Let $x_1, \ldots, x_5$ be covariates

$$M = \cup_{k=1}^{3} M_k,$$

where

$$\begin{aligned} M_1 &= \{\{x_1, x_3\}, \{x_2\}\}, \\ M_2 &= \{\{x_2, x_3, x_4\}, \{x_3, x_4\}\}, \\ M_3 &= \{\{x_3, x_5\}, \{x_3\}\}, \end{aligned}$$

$$\begin{aligned} m_{2,inf} &= \{x_3, x_4\} \\ m_{2,sup} &= \{x_2, x_3, x_4\} \end{aligned}$$

Branch and bound for model selection

Reminder:

For the M_k, let

$m_{k,inf}$ be the largest model contained in every model in M_k

$m_{k,sup}$ be a smallest model that contains every model in M_k

Then, $\forall m \in M_k$

$$F(m) \geq \hat{R}(\hat{\beta}_{m_{k,\text{sup}}}) + 2|m_{k,\text{inf}}| = L_k$$

$$F(m) \leq \hat{R}(\hat{\beta}_{m_{k,\text{inf}}}) + 2|m_{k,\text{sup}}| = U_k$$

Branch and bound for model selection: An algorithm

1. Define a global variable $b = F(m)$ for any $m \in M$

   (As an aside, every time $F(m)$ is computed, update b if $F(m) < b$)
2. Partition $M = \{M_1, \ldots, M_K\}^\sharp$
3. For each k, if $L_k > b$, eliminate the branch M_k
4. Else, recurse and return to 2., substituting M_k for M