1 Постановка

Преамбула. $H = \{h_0, h_1, ..., h_r\}$ — множество базисных функций с арностями $a_0, a_1, ..., a_r \geq 0$. Предполагается, что h_0 является корнем дерева с арностью $a_0 = 1$.

$\Gamma = (V, E)$ — дерево с вершинами $v_i \in H$, выбранными из H с возвращением. Количество исходящих ребер каждой вершины v_i определяется значением соответствующей арности a_i.

$p(h_i, h_j)$ — вероятность того, что базисная функция h_j является дочерью базисной функции h_i.

Независимость от посторонних вершин: вероятность наблюдать дерево $P(\Gamma) = \prod\limits_{(i,j)\in E} p(v_i, v_j)$.

Амбула. Для объекта x на H определен неизвестный набор вероятностей $p_x(h_i, h_j) = f_{ij}(x)$. Необходимо для объекта x найти дерево $\hat{\Gamma}$, максимизирующее его вероятностную конфигурацию

$$\hat{\Gamma} = \arg\max_{(V,E)} \prod_{(i,j)\in E} p(v_i, v_j). \tag{1}$$

2 Решение

Оценка вероятностей. Задана выборка $(x_k, \Gamma_k)_{k=1}^m$. Оценим функции $\hat{f}_{ij}(x) = p_x(h_i, h_j)$ с использованием значений этих функций на заданной выборке: $f_{ij}(x_k) = [(h_i, h_j) \in E_k]$.

Построение дерева. По найденной матрице $P_x : P_x(i,j) = \hat{p}_x(h_i, h_j)$ построим дерево $\hat{\Gamma}$, максимизирующее вероятностную конфигурацию (1). Предлагается жадный алгоритм. Введем два дополнительных понятия.

Вершину $v_i \in H$ назовем *открытой*, если в текущей конфигурации (V, E) количество ее дочерей меньше, чем арность вершины a_i.

Множество $H_0 \subset H$ функций нулевой арности назовем *множеством свободных переменных*. Элементы множества H_0, и только они, являются листьями дерева Γ.

1. $V = \{h_0\}, E = \emptyset, s_{pred} = 0$.

2. Пока в (V, E) есть открытые вершины v^{op}: (3-5)

3. Закроем v^{op} оптимальными свободными переменными $v^0 \in H_0$:

   $$v^0 \to V, \quad e(v^{op}, v^0) \to E, \quad v^0 = \arg\max_{v\in H_0} \hat{p}_x(v^{op}, v) \quad \text{для всех } v^{op}.$$

   Обозначим текущую конфигурацию за Γ_0.

4. Присвоим $s_{new} = P(\Gamma_0)$. Если $s_{new} < s_{pred}$, то выход, вернуть конфигурацию, соответствующую s_{pred}. Иначе присвоим $s_{pred} = s_{new}$.

5. Добавим в открытую конфигурацию (V, E) оптимальную вершину из H:
$$v^* \to V, \quad e(v^{op}, v^*) \to E, \quad v^* = \arg \max_{v \in H, v^{op}} \hat{p}_x(v^{op}, v).$$