Multivariate Dyadic Regression Trees for Sparse Learning Problems
Appendix

Han Liu and **Xi Chen**
School of Computer Science, Carnegie Mellon University
Pittsburgh, PA 15213

1 Pseudocode for Greedy Tree Learning Algorithm

Algorithm 1 Tree Growing

Input: $\{\mathbf{x}^i, \mathbf{y}^i\}_{i=1}^n$, m, $n_{\max}$, L, *deterministic*

```
Build the initial tree T with a single root node r containing all the data points and set A(r) ← ∅
while ∃t ∈ term(T) such that n(t) > n_max or |A(t)| < d do
  if |A(t)| < d then
    for all dimension k ∉ A(t) do
      calculate ΔR̂_1^m(t, k)
    end for
  end if
  if n(t) > n_max then
    for all dimension k ∈ A(t) do
      if sl(t^(k)) ≥ 2^(−L+1) then
        calculate ΔR̂_2^m(t, k)
      end if
    end for
  end if
  if deterministic then
    (a*, k*) = argmax_{a∈{1,2}, k∈{1...d}} ΔR̂_a^m(t, k)
  else
    Normalize ΔR̂ to 1. Draw the sample (a*, k*) from the multinomial(1, ΔR̂)
  end if
  if a* = 1 then
    Dyadically split the cell represented by node t perpendicular to dimension k* and update T
  else
    A(t) ← A(t) ∪ {k}
  end if
end while
```

Output: Tree T

Note the boolean variable *deterministic* indicates whether the procedure is purely greedy or randomized.

Algorithm 2 Cost Complexity Pruning

Input: Tree T, parameter λ for calculating $\widehat{C}(T)$

$i \leftarrow 1, T_1 \leftarrow T$

while T_i has more than one node OR T_i only has the root node r with $\mathcal{A}(r) \neq \emptyset$ **do**

$$T^{(1)} \leftarrow \underset{t_L, t_R \in \text{term}(T_i)}{\text{argmin}} \widehat{C}(\text{ Tree obtained by merging } t_L, t_R \text{ in } T_i)$$

$$T^{(2)} \leftarrow \underset{t \in \text{term}(T_i), k \in \mathcal{A}(t) \setminus \mathcal{A}(\text{par}(t))}{\text{argmin}} \widehat{C}(\text{ Tree obtained by removing the dimension } k \text{ from } \mathcal{A}(t))$$

$T_{i+1} \leftarrow \underset{T^{(l)} \in \{1,2\}}{\text{argmin}} \widehat{C}(T^{(l)})$

$i \leftarrow i + 1$

end while

$i^* \leftarrow \underset{i}{\text{argmin}} \widehat{C}(T_i)$

Output: Optimal Tree T_{i^*}