Piecewise linear concave model. In this section, we propose a nonparametric method for multivariate regression subject to concavity of the value function. We introduce concave/convex adaptive regression tree (CARET) which adaptively partitions the state space and fit a local linear regression on each partition. Therefore, the resulted model unlike the piecewise regression model is a concave/convex continuous function. The convexity/concavity property is important in particular in operation research literature since it decreases the computational time and increases the chance of finding a global optima. Given B subsets of linear model with coefficients $(\alpha, \boldsymbol{\beta})^k, \forall k \in \{1, \cdots, K\}$

$$\hat{Y} = \min_{k \in \{1, \cdots, K\}} \alpha^k + \boldsymbol{\beta}^k \boldsymbol{X}$$

The partitioning happens in two stages. First, we split along a state variable in $\boldsymbol{X}$ to decrease the global least square error and increase the number of partitions, K, and fit a linear regression on each subset to construct a hyperplane. Then, we reallocate the each point to the subset, the hyperplace of which is dominant. This refit step places the hyperplanes in a closer alignment with observations in the dataset. The process continues until the reallocation of the observations does not create a change in the partitions or all subsets have a minimal number of observations. Suppose our data consists of p inputs and a response for each N observations.

Algorithm 2 adaptively partition the observations and fit hyperplanes on the resulted subsets. To ensure the concavity/convexity of the estimation, it refits the observations to the dominant hyperplane for each observation. We start the model with $K = 1$, and split the data on the dimension l and threshold s into two subsets to minimize the square loss. In the next step we run a linear regression on each subset to find the corresponding coefficients $(\alpha, \boldsymbol{\beta})$. A split is accepted by the algorithm if the number of observations in a subset is more than predefined minimum value n_{min}. Otherwise, we reject the split and move to the next subset. Once the algorithm evaluate the cuts on all subsets, it refits using the partition induced by hyperplanes. The model stops either if there are not enough observations within each subset $|J_{k,iter}| < n_{min}$, or if there is no change in the subsets $J_{k,iter+1} = J_{k,iter}, \forall k \in K$.

Algorithm 2: CARET

Result: $(\alpha, \boldsymbol{\beta})^k$

Set $K = 1$, n_{min}, $\boldsymbol{X} \in J_1$, $iter = 1$, and Flag = True

while *Flag = True* **do**

  $K_{new} = K$;

  $counter = 0$;

  for $k \in K$ **do**

    $\text{argmin}_{l,s} \left[\min_{c_1} \sum_{x_i^k \in J'_{k,iter}(l,s)} (y_i^k - c_1)^2 + \min_{c_2} \sum_{x_i^k \in J''_{k,iter}(l,s)} (y_i^k - c_2)^2 \right]$,

    where $(i, l) \in (N, p), x_i^k \in J_{k,iter}$, $J'_{k,iter}(l, s) = \{\boldsymbol{X} | \boldsymbol{X}_l \leqslant s\}$,

    $J''k, iter(l, s) = \{X | X_l > s\}$, and $c_h = ave(y_i | \boldsymbol{X}_i \in J'_h(l, s))$;

    if $|J'_{k,iter}| > n_{min}$ *and* $|J''_{k,iter}| > n_{min}$ **then**

      $K_{new} = K_{new} + 1$;

      $(\alpha', \boldsymbol{\beta'})^{counter+1} = (\boldsymbol{X}^\intercal_{\boldsymbol{J'_k}} \boldsymbol{X}_{\boldsymbol{J'_k}})^{-1} \boldsymbol{X}^\intercal_{\boldsymbol{J'_k}} y_{J'_k}$;

      $(\alpha', \boldsymbol{\beta'})^{counter+1} = (\boldsymbol{X}^\intercal_{\boldsymbol{J''_k}} \boldsymbol{X}_{\boldsymbol{J''_k}})^{-1} \boldsymbol{X}^\intercal_{\boldsymbol{J''_k}} y_{J''_k}$;

    else

      $(\alpha', \boldsymbol{\beta'})^{counter+1} = (\alpha, \boldsymbol{\beta})^k$;

  $(\alpha, \boldsymbol{\beta}) = (\alpha', \boldsymbol{\beta'})$;

  $K = K_{new}$;

  $J_{iter+1} = \{J'_{iter}, J''_{iter}\}$;

  $J_{k,iter+1} = \{\boldsymbol{X_i} : \alpha^k + \boldsymbol{\beta}^k \boldsymbol{X}_i \leqslant \alpha^{k\prime} + \boldsymbol{\beta}^{k\prime} \boldsymbol{X}_i, k' \neq k\}$;

  if $|J_{k,iter}| < n_{min}, \forall k \in K$ *or* $J_{k,iter+1} = J_{k,iter}, \forall k \in K$ **then**

    Flag = False;

  else

    $iter + = 1$

$\min_{k \in \{1, \cdots, K\}} \alpha^k + \boldsymbol{\beta}^k \boldsymbol{X}$