consists of an intercept parameter \((\beta_0)\) and the weighted (by \(\beta_m\)) sum of one or more basis functions \(h_m(X)\), of the kind illustrated earlier. You can also think of this model as “selecting” a weighted sum of basis functions from the set of (a large number of) basis functions that span all values of each predictor (i.e., that set would consist of one basis function and parameter \(t\), for each distinct value for each predictor variable). The MARSplines algorithm then searches over the space of all inputs and predictor values (knot locations \(t\)) as well as interactions between variables. During this search, an increasingly larger number of basis functions is added to the model (selected from the set of possible basis functions), to maximize an overall least squares goodness-of-fit criterion. As a result of these operations, MARSplines automatically determines the most important independent variables as well as the most significant interactions among them. The details of this algorithm are further described in Hastie et al. (2001).

**Categorical Predictors**

MARSplines is well suited for tasks involving categorical predictors variables. Different basis functions are computed for each distinct value for each predictor, and the usual techniques for handling categorical variables are applied. Therefore, categorical variables (with class codes rather than continuous or ordered data values) can be accommodated by this algorithm without requiring any further modifications.

**Multiple Dependent (Outcome) Variables**

The MARSplines algorithm can be applied to multiple dependent (outcome) variables, whether continuous or categorical. When the dependent variables are continuous, the algorithm will treat the task as regression; otherwise, as a classification problem. When the outputs are multiple, the algorithm will determine a common set of basis functions in the predictors but estimate different coefficients for each dependent variable. This method of treating multiple outcome variables is not unlike some neural network architectures, where multiple outcome variables can be predicted from common neurons and hidden layers; in the case of MARSplines, multiple outcome variables are predicted from common basis functions, with different coefficients.

**MARSplines and Classification Problems**

Because MARSplines can handle multiple dependent variables, it is easy to apply the algorithm to classification problems as well. First, it will code the classes in the categorical response variable into multiple indicator variables (e.g., \(1 = \) observation belongs to class \(k\), \(0 = \) observation does not belong to class \(k\)); then MARSplines will fit a model and compute predicted (continuous) values or scores; and finally, for prediction, it will assign each case to the class for which the highest score is predicted (see also Hastie et al., 2001, for a description of this procedure).