HW-10

Tamoghna

Q1 <u>True</u>

When we use L+1, on say R_m no. of regions, dividing the 2-D feature space. Using such a combination of divisions, where we have a boundary parallel to each axis, it is possible to generate any region and assign any possible class to that region.

Axis parallel splits can divide the region into R_m or L+1 regions. The number of such splits will define the regions (basis functions) and the weights specify the response value in each region. The regions we get from such splits will surely be complex and therefore will satisfy the orientation.

Thus, hypothesis set H truly describes hypothesis set for this CART algorithm.

Q2. (a) Forward stagewise additive modeling:

$$f_m(x) = f_{m-1}(x) + \beta_m \phi(x; y_m) \quad \text{— ①}$$

(16.34)

Additive Basis Function Model:

$$f(x) = w_0 + \sum_{m=1}^{M} w_m \phi_m(x) \quad \text{— ②}$$

(16.3)

Using ①,

$$f_1(x) = f_0(x) + \beta_1 \phi(x; y_1)$$

$$f_2(x) = f_1(x) + \beta_2 \phi(x; y_2)$$

$$= f_0(x) + \beta_1 \phi(x; y_1) + \beta_2 \phi(x; y_2)$$

So,

$$f_m(x) = f_{m-1}(x) + \beta_m \phi(x; y_m)$$

PTO

becomes,

$$f_m(x) = f_0(x) + \beta_1 \phi(x; y_1) + \beta_2 \phi(x; y_2) + \cdots + \beta_M \phi(x; y_m)$$

using ②,

$$f(x) = w_0 + w_1 \phi_1(x) + w_2 \phi_2(x) + \cdots + w_m \phi_m(x)$$

Comparing, we get,

$$\boxed{\begin{aligned} w_0 &= f_0(x) = \underset{y}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, f(x_i; y)) \quad \text{(eqn 16.32)} \\ w_m &= \beta_m \\ \phi_m &= \phi(x; y_m) \end{aligned}}$$

(b) Shrinkage used in the stagewise additive modeling:

$$f_m(x) = f_{m-1}(x) + \nu \beta_m \phi(x; y_m) \quad \text{—— ①}$$

(Eq 16.35)

Additive basis function model:

$$f(x) = w_0 + \sum_{m=1}^{M} w_m \phi_m(x) \quad \text{—— ②}$$

(Eq 16.3)

Using ①,

$$f_1(x) = f_0(x) + \nu \beta_1 \phi(x; y_1)$$

$$f_2(x) = f_1(x) + \nu \beta_2 \phi(x; y_2)$$

$$= f_0(x) + \nu \beta_1 \phi(x; y_1) + \nu \beta_2 \phi(x; y_2)$$

$$\vdots$$

$$f_m(x) = f_{m-1}(x) + \nu \beta_m \phi(x, y_m)$$

$$= f_0(x) + \nu \beta_1 \phi(x; y_1) + \cdots$$

$$+ \nu \beta_M \phi(x; y_M)$$

Using ②,

$$f(x) = w_0 + w_1 \phi_1(x) + w_2 \phi_2(x) + \cdots + w_m \phi_m(x)$$

Comparing, we get,

$$\boxed{\begin{aligned} w_0 &= f_0(x) = \operatorname{argmin} \sum_{i=1}^{N} L(y_i, f(x_i; y)) \quad \text{(eqn 16.32)} \\ w_m &= \nu \beta_m \\ \phi_m &= \phi(x; y_m) \end{aligned}}$$