Theory of Generalization

Starfly
starfly3119@gmail.com

Beihang University — February 18, 2020

1 Restriction of Break Point

what 'must be true' when minimum beak point k = 2

- N = 1: every $m_H(N) = 2$ by definition
- N = 2: every $m_H(N) < 4$ by definition
  (so maximum possible = 3)

maximum possible $m_H(N)$ when $N = 3$ and $k = 2$?

maximum possible so far: 4 dichotomies

- N = 3: maximum possible = $4 << 2^3$

Break point k restricts maximum possible $m_H(N)$ a lot for $N > k$

idea:

$$\begin{aligned} m_H(N) &\leq \textbf{maximum possible } m_H(N) \textbf{ given k} \\ &\leq poly(N) \end{aligned} \tag{1}$$

2 Bounding Function: Basic Cases

2.1 Definition

bounding function $B(N, k)$: maximum possible $m_H(N)$ when break point = k

- combinatorial quantity:
  maximum number of length-N vectors with (o,x)
  while 'no shatter' any length-k subvectors
- irrelevant of the details of H
  e.g. $B(N, 3)$ bounds both
  - positive intervals (k = 3)
  - 1D perceptrons (k = 3)

new goal: $B(N, k) \leq poly(N)$?

2.2 Table of Bounding Function

3 Bounding Functions: Inductive Cases

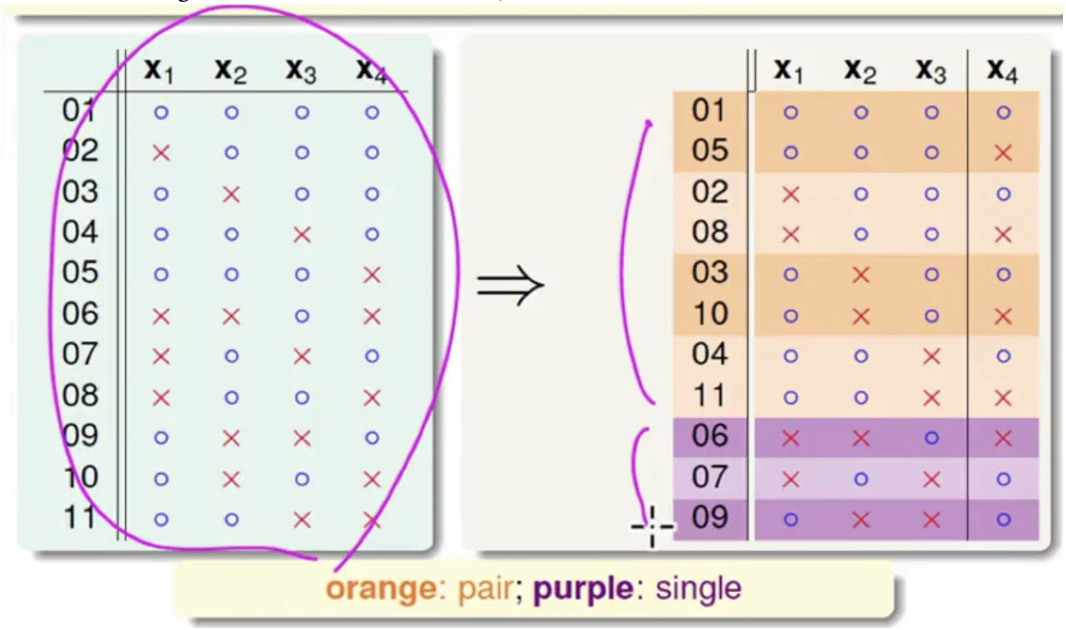
3.1 Estimating $B(4,3)$

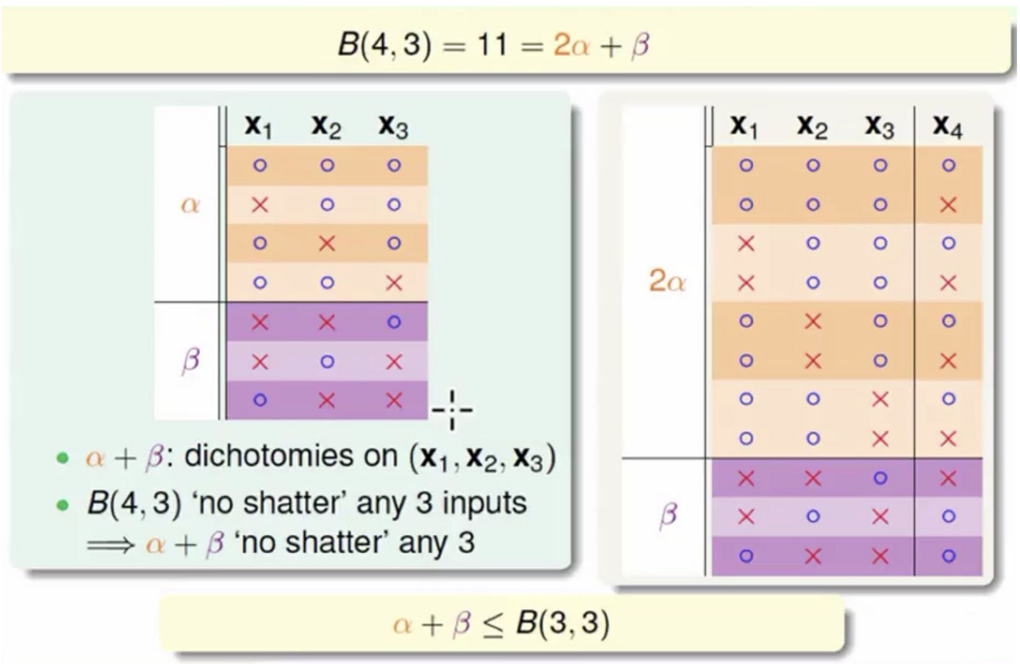
Motivation

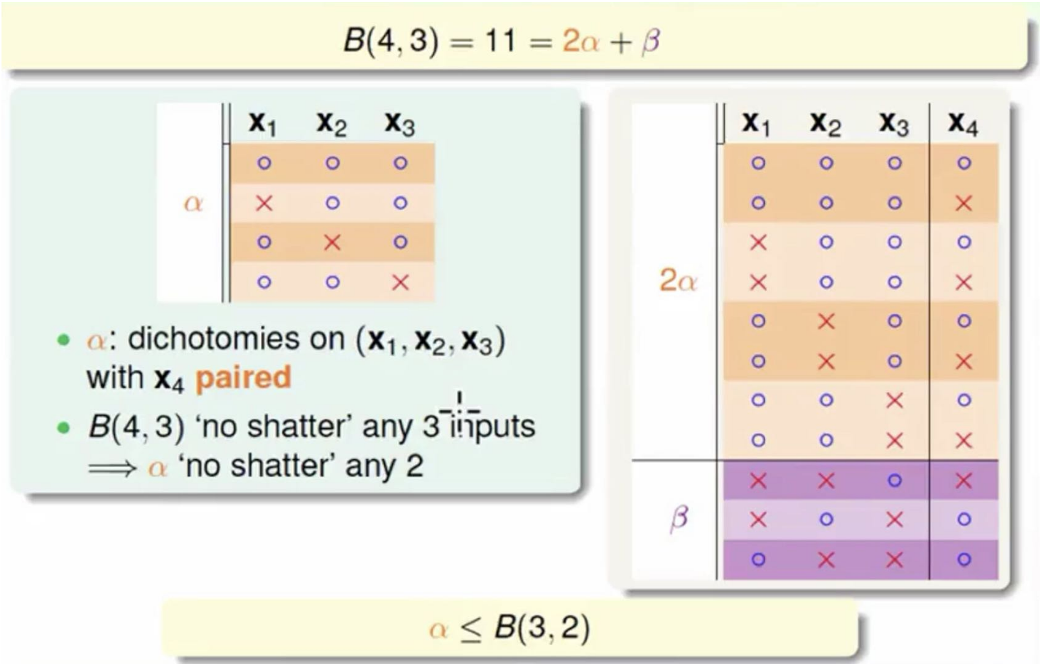
- $B(4,3)$ shall be related to $B(3,?)$
  'adding' one point from $B(3,?)$

next: reduce $B(4,3)$ to $B(3,?)$

after checking all 2^16 sets of dichotomies, the winner is...







$$
\begin{aligned}
B(N,k) &= 2\alpha + \beta \\
\alpha + \beta &\leq B(N-1,k) \\
\alpha &\leq B(N-1,k-1)
\end{aligned}
\tag{2}
$$

$$B(N,k) \leq B(N-1,k) + B(N-1,k-1) \tag{3}$$

3.2 The Theorem

$$B(N,k) \leq \sum_{i=0}^{k-1} C_N^i N^i \tag{4}$$

- simple induction using boundary and inductive formula
- for fixed k, $B(N,k)$ upper bounded by poly(N)
- actually $\leq$ can be $=$ $m_H(N)$ is poly(N) if break point exists

4 A Pictorial Proof

Bad Bound for General H,

actually, when N large enough,

$$P[\exists h \in H, s.t. |E_{in}(h) - E_{out}(h)| > \epsilon] \leq 2 \cdot 2m_H(2N) \cdot exp(-2 \cdot \frac{1}{16}\epsilon^2 N) \tag{5}$$