Review of Lecture 4

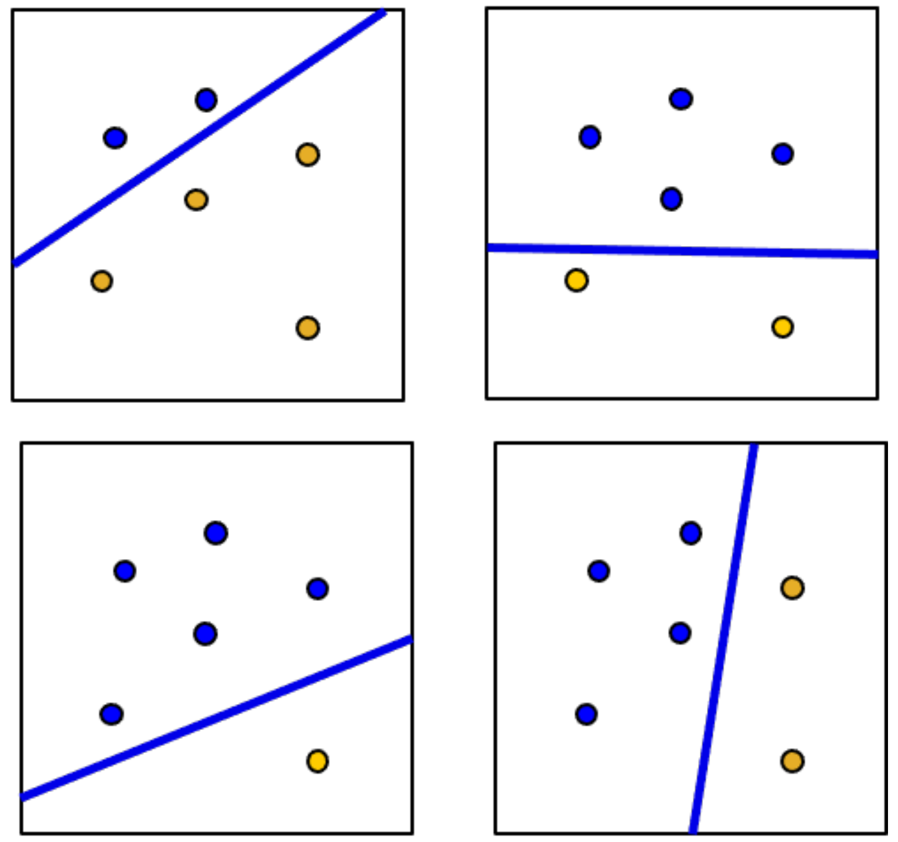
Instead of

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \leq 2 \quad M \quad e^{-2N\epsilon^2}$$

seek to replace M with **growth function**

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \overset{?}{\leq} 2 \quad m_{\mathcal{H}}(N) \quad e^{-2N\epsilon^2}$$

Dichotomies



Review of Lecture 4

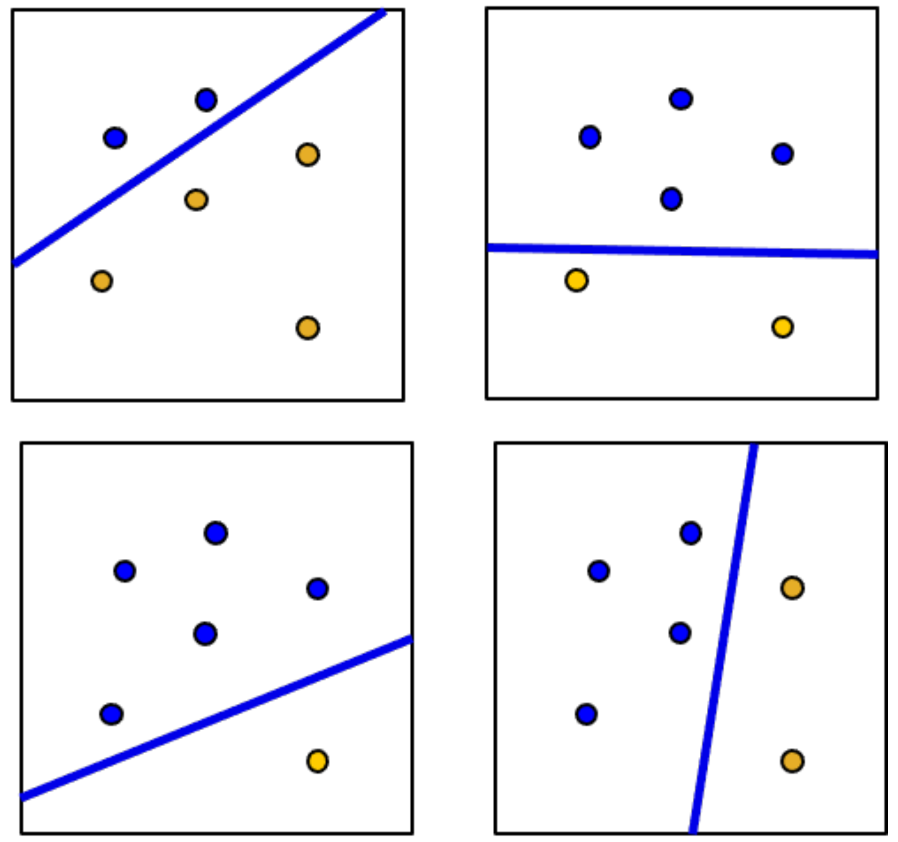
Instead of

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \leq 2 \quad {\color{red}M} \quad e^{-2N\epsilon^2}$$

seek to replace M with **growth function**

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \overset{?}{\leq} 2 \quad {\color{red}m_{\mathcal{H}}(N)} \quad e^{-2N\epsilon^2}$$

Dichotomies



$${\color{red}m_{\mathcal{H}}(N)} = \max_{\mathbf{x}_1,\ldots,\mathbf{x}_N} |\mathcal{H}(\mathbf{x}_1,\ldots,\mathbf{x}_N)|$$

Review of Lecture 4

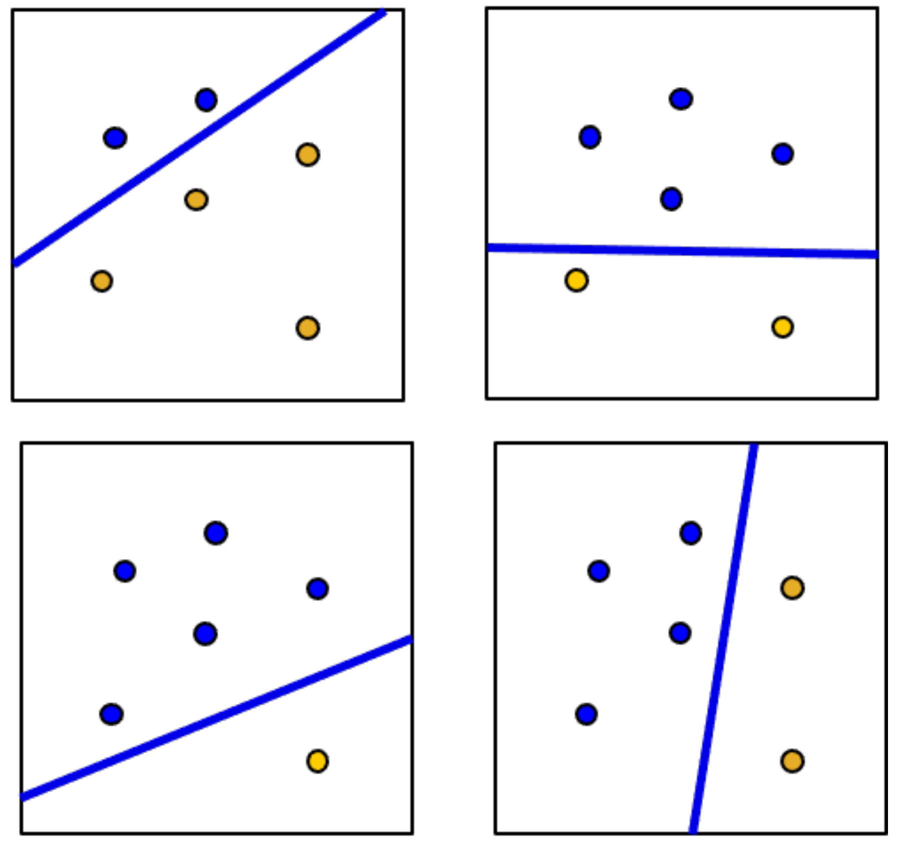
Instead of

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \leq 2 \quad M \quad e^{-2N\epsilon^2}$$

seek to replace M with **growth function**

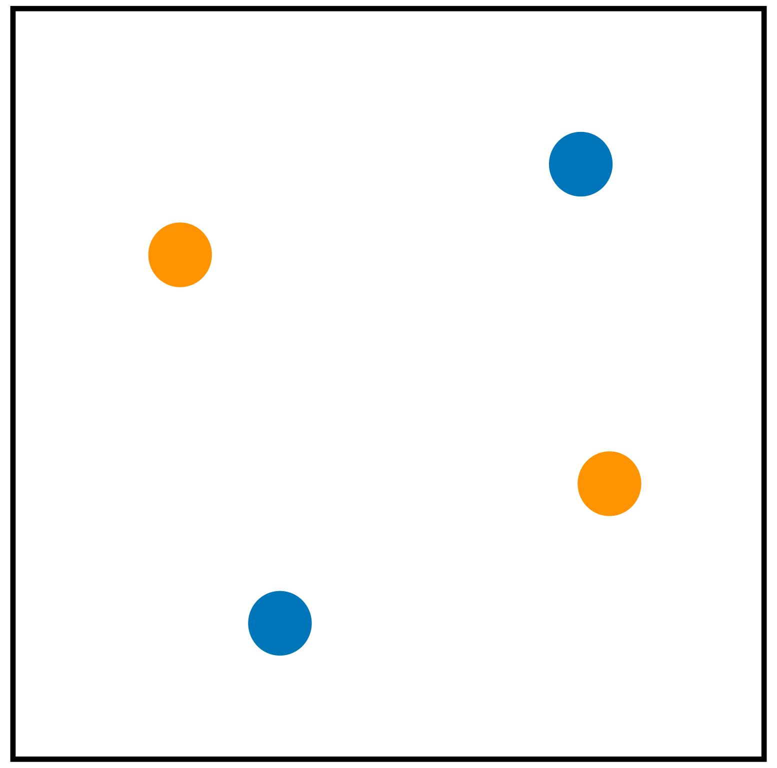
$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \overset{?}{\leq} 2 \quad m_{\mathcal{H}}(N) \quad e^{-2N\epsilon^2}$$

Dichotomies



$$m_{\mathcal{H}}(N) = \max_{\mathbf{x}_1,\ldots,\mathbf{x}_N} |\mathcal{H}(\mathbf{x}_1,\ldots,\mathbf{x}_N)|$$

- Break point $k = 4$



Review of Lecture 4

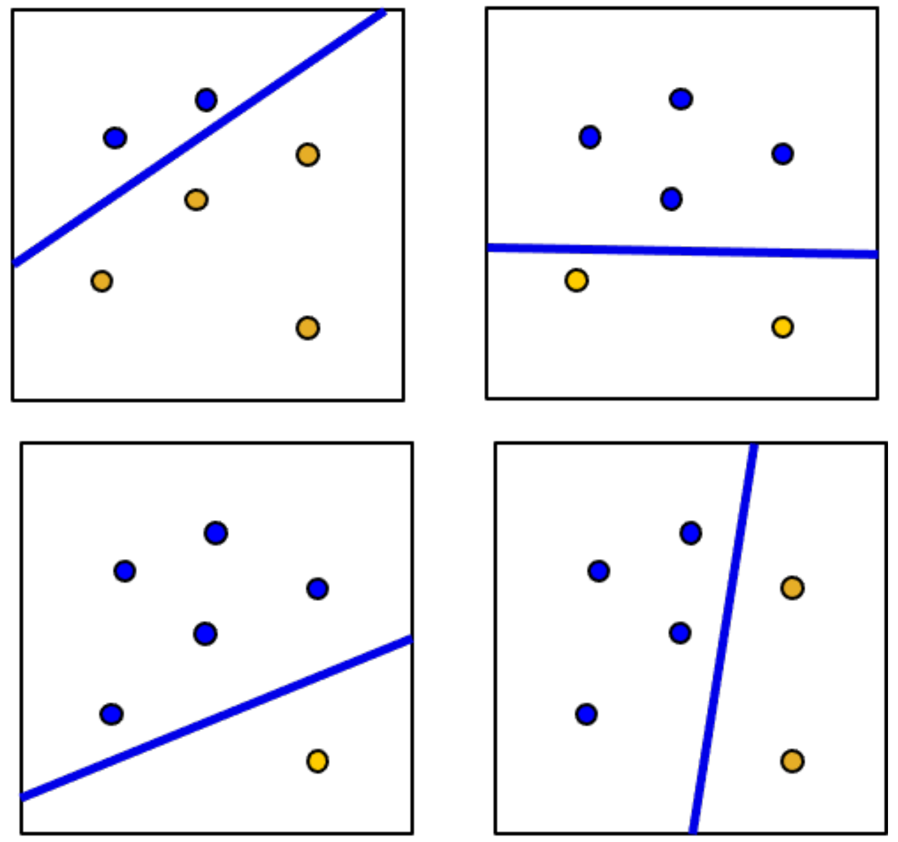
Instead of

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \leq 2 \quad {\color{red}M} \quad e^{-2N\epsilon^2}$$

seek to replace M with **growth function**

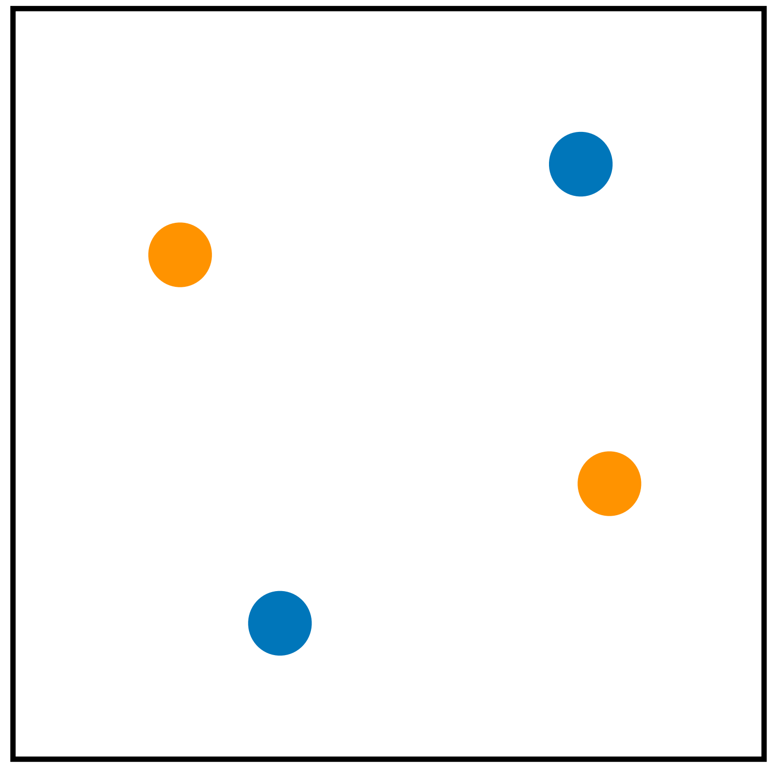
$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \overset{?}{\leq} 2 \quad {\color{red}m_{\mathcal{H}}(N)} \quad e^{-2N\epsilon^2}$$

Dichotomies



$${\color{red}m_{\mathcal{H}}(N)} = \max_{\mathbf{x}_1,\ldots,\mathbf{x}_N} |\mathcal{H}(\mathbf{x}_1, \ldots, \mathbf{x}_N)|$$

- Break point $k = 4$



- Maximum # of dichotomies

x_1	x_2	x_3
○	○	○
○	○	●
○	●	○
●	○	○

$k = 2$

ECE 6254
Statistical Machine Learning

Professor: Amirali Aghazadeh
Office: Coda S1209
Georgia Institute of Technology

Lecture 5: Theory of Generalization II

Outline

- Prove that $m_{\mathcal{H}}(N)$ is polynomial

- Prove that $m_{\mathcal{H}}(N)$ can replace M

Why do we like polynomial growth so much? (and dislike exponential)

A) Union Bound $\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \leq 2Me^{-2N\epsilon^2}$

B) With probability at least $1-\delta$: $R(h^*) \leq \hat{R}_N(h^*) + \sqrt{\frac{1}{2N}\ln\frac{2M}{\delta}}$

Assume *M* grows

1) exponential with N
2) polynomial with N

what happens when N goes to infinity?

Bounding $m_{\mathcal{H}}(N)$

To show: $m_{\mathcal{H}}(N)$ is polynomial

We show: $m_{\mathcal{H}}(N) \leq \cdots \leq$ a polynomial

Bounding $m_{\mathcal{H}}(N)$

To show: $m_{\mathcal{H}}(N)$ is polynomial

We show: $m_{\mathcal{H}}(N) \leq \cdots \leq$ a polynomial

Key quantity: $B(N, k)$

Maximum number of dichotomies on N points, with break point k

Recall the Puzzle from the last lecture

$B(N, k)$ for $N = 3, k = 2$

x_1	x_2	x_3
○	○	○
○	○	●
○	●	○
~~○~~	~~●~~	~~●~~
●	○	○
~~●~~	~~○~~	~~●~~
~~●~~	~~●~~	~~○~~
~~●~~	~~●~~	~~●~~

x_1	x_2	x_3
0	0	0
0	0	1
0	1	0
~~0~~	~~1~~	~~1~~
1	0	0
~~1~~	~~0~~	~~1~~
~~1~~	~~1~~	~~0~~
~~1~~	~~1~~	~~1~~

Recall the Puzzle from the last lecture

$B(N,k)$ for $N=3, k=2$

$S_1 \triangleq$ dichotomies appear only once on 1st N-1 columns

$S_2 \triangleq$ dichotomies appear twice on 1st N-1 columns

x_1	x_2	x_3	
0	0	0	S_2^-
0	0	1	S_2^+
0	1	0	S_1
~~0~~	~~1~~	~~1~~	
1	0	0	S_1
~~1~~	~~0~~	~~1~~	
~~1~~	~~1~~	~~0~~	
~~1~~	~~1~~	~~1~~	

Recursive bound on $B(N, k)$

Let's expand this to larger N

x_1	x_2	$\dots$	x_{N-1}	x_N

Recursive bound on $B(N,k)$

Let's expand this to larger N

x_1	x_2	$\dots$	x_{N-1}	x_N
0	0	$\dots$	0	0

Recursive bound on $B(N,k)$

Let's expand this to larger N

		x_1	x_2	$\dots$	x_{N-1}	x_N
		0	0	$\dots$	0	0
		1	0	$\dots$	0	1
S_1		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	1	1
		1	0	$\dots$	1	0

Recursive bound on $B(N,k)$

Let's expand this to larger N

$$B(N,k) = \alpha \ +$$

	# of rows	x_1	x_2	$\ldots$	x_{N-1}	x_N
		0	0	$\ldots$	0	0
		1	0	$\ldots$	0	1
S_1	α	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	1	1
		1	0	$\ldots$	1	0

Recursive bound on $B(N,k)$

Let's expand this to larger N

$$B(N,k) = \alpha \ +$$

	# of rows	x_1	x_2	$\ldots$	x_{N-1}	x_N
S_1	α	0	0	$\ldots$	0	0
		1	0	$\ldots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	1	1
		1	0	$\ldots$	1	0
S_2^-		0	1	$\ldots$	0	0
S_2^+		0	1	$\ldots$	0	1

Recursive bound on $B(N, k)$

Let's expand this to larger N

$$B(N, k) = \alpha \ +$$

	# of rows	x_1	x_2	$\dots$	x_{N-1}	x_N
S_1	α	0	0	$\dots$	0	0
		1	0	$\dots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	1	1
		1	0	$\dots$	1	0
S_2^-		0	1	$\dots$	0	0
		1	1	$\dots$	0	0
S_2^+		0	1	$\dots$	0	1
		1	1	$\dots$	0	1

Recursive bound on $B(N,k)$

Let's expand this to larger N

$$B(N,k) = \alpha + 2\beta$$

	# of rows	x_1	x_2	$\ldots$	x_{N-1}	x_N
S_1	α	0	0	$\ldots$	0	0
		1	0	$\ldots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	1	1
		1	0	$\ldots$	1	0
S_2^-	β	0	1	$\ldots$	0	0
		1	1	$\ldots$	0	0
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	0	0
		1	1	$\ldots$	1	0
S_2^+	β	0	1	$\ldots$	0	1
		1	1	$\ldots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	0	1
		1	1	$\ldots$	1	1

Estimating α and β

Focus on $x_1, \ldots, x_{N-1}$ columns

	# of rows	x_1	x_2	$\ldots$	x_{N-1}	x_N
		0	0	$\ldots$	0	0
		1	0	$\ldots$	0	1
S_1	α	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	1	1
		1	0	$\ldots$	1	0
		0	1	$\ldots$	0	0
		1	1	$\ldots$	0	0
S_2^-	β	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	0	0
		1	1	$\ldots$	1	0
		0	1	$\ldots$	0	1
		1	1	$\ldots$	0	1
S_2^+	β	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	0	1
		1	1	$\ldots$	1	1

Estimating α and β

Focus on $x_1, \ldots, x_{N-1}$ columns

$$\alpha + \beta \leq B(N-1, k)$$

	# of rows	x_1	x_2	$\ldots$	x_{N-1}	x_N
S_1	α	0	0	$\ldots$	0	0
		1	0	$\ldots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	1	1
		1	0	$\ldots$	1	0
S_2^-	β	0	1	$\ldots$	0	0
		1	1	$\ldots$	0	0
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	0	0
		1	1	$\ldots$	1	0
S_2^+	β	0	1	$\ldots$	0	1
		1	1	$\ldots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\ldots$	0	1
		1	1	$\ldots$	1	1

Estimating β by itself

Focus on $S_2 = S_2^+ \cup S_2^-$ rows

	# of rows	x_1	x_2	$\dots$	x_{N-1}	x_N
S_1	α	0	0	$\dots$	0	0
		1	0	$\dots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	1	1
		1	0	$\dots$	1	0
S_2^-	β	0	1	$\dots$	0	0
		1	1	$\dots$	0	0
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	0	0
		1	1	$\dots$	1	0
S_2^+	β	0	1	$\dots$	0	1
		1	1	$\dots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	0	1
		1	1	$\dots$	1	1

Estimating β by itself

Focus on $S_2 = S_2^+ \cup S_2^-$ rows

$$\beta \leq B(N-1, k-1)$$

	# of rows	x_1	x_2	$\dots$	x_{N-1}	x_N
S_1	α	0	0	$\dots$	0	0
		1	0	$\dots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	1	1
		1	0	$\dots$	1	0
S_2^-	β	0	1	$\dots$	0	0
		1	1	$\dots$	0	0
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	0	0
		1	1	$\dots$	1	0
S_2^+	β	0	1	$\dots$	0	1
		1	1	$\dots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	0	1
		1	1	$\dots$	1	1

Putting it all together

$$B(N,k) = \alpha \ + 2\beta$$

$$\alpha + \beta \leq B(N-1,k)$$

$$\beta \leq B(N-1,k-1)$$

$$B(N,k) \leq$$

$$B(N-1,k) + B(N-1,k-1)$$

	# of rows	x_1	x_2	$\dots$	x_{N-1}	x_N
S_1	α	0	0	$\dots$	0	0
		1	0	$\dots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	1	1
		1	0	$\dots$	1	0
S_2^-	β	0	1	$\dots$	0	0
		1	1	$\dots$	0	0
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	0	0
		1	1	$\dots$	1	0
S_2^+	β	0	1	$\dots$	0	1
		1	1	$\dots$	0	1
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
		0	1	$\dots$	0	1
		1	1	$\dots$	1	1

Numerical computation of $B(N,k)$ bound

$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$

		k						
		1	2	3	4	5	6	..
	1							
	2							
	3							
N	4							
	5							
	6							
	$\vdots$							

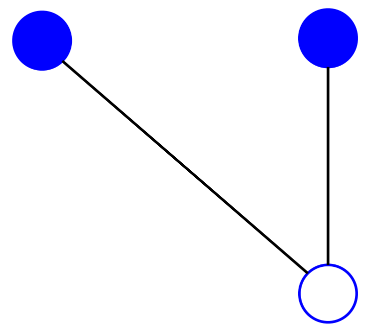
Numerical computation of $B(N,k)$ bound

$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$

		k						
		1	2	3	4	5	6	..
	1	1						
	2	1						
	3	1						
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

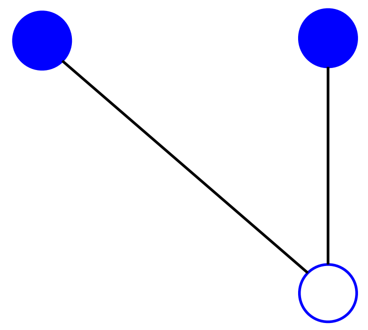
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



		k						
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1						
	3	1						
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

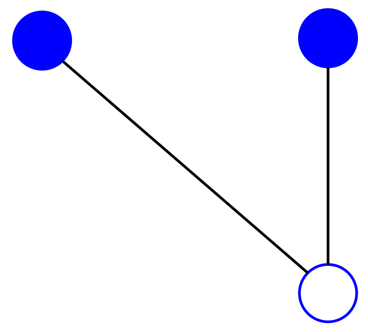
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



		k						
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3					
	3	1						
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

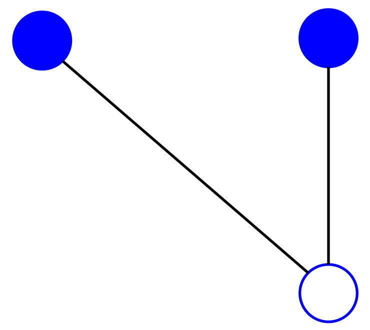
$$B(N,k) \le B(N-1,k) + B(N-1,k-1)$$



		k						
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3	4				
	3	1						
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

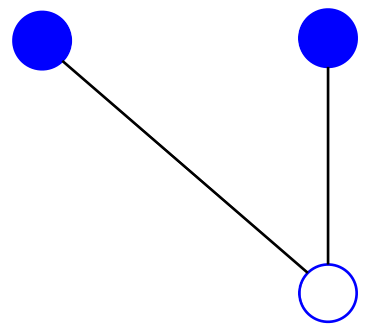
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



					k			
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3	4	4	4	4	..
	3	1						
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

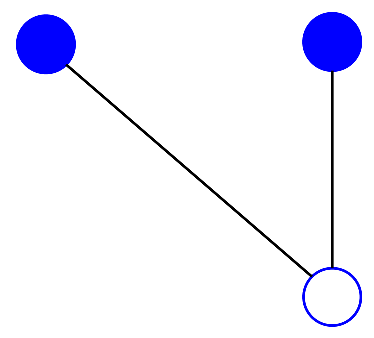
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



					k			
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3	4	4	4	4	..
	3	1	4					
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

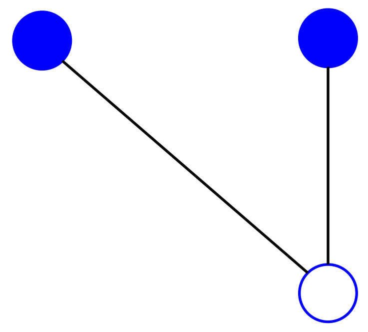
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



		k						
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3	4	4	4	4	..
	3	1	4	7				
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

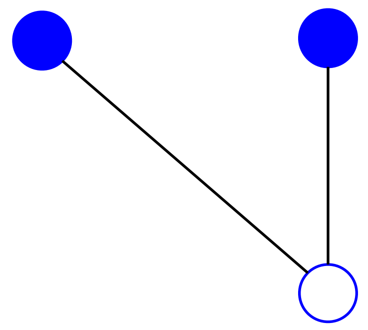
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



		k						
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3	4	4	4	4	..
	3	1	4	7	8			
N	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

Numerical computation of $B(N,k)$ bound

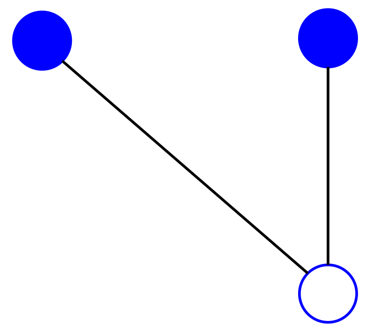
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



		k						
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3	4	4	4	4	..
	3	1	4	7	8	8	8	..
N	4	1	5	..	..			
	5	1	6					
	6	1	7					
	$\vdots$	$\vdots$	$\vdots$					

Numerical computation of $B(N,k)$ bound

$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



		k						
		1	2	3	4	5	6	..
	1	1	2	2	2	2	2	..
	2	1	3	4	4	4	4	..
	3	1	4	7	8	8	8	..
N	4	1	5	..	..			
	5	1	6					
	6	1	7					
	:	:	:					

Numerical computation of $B(N,k)$ bound

$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$

Theorem:

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

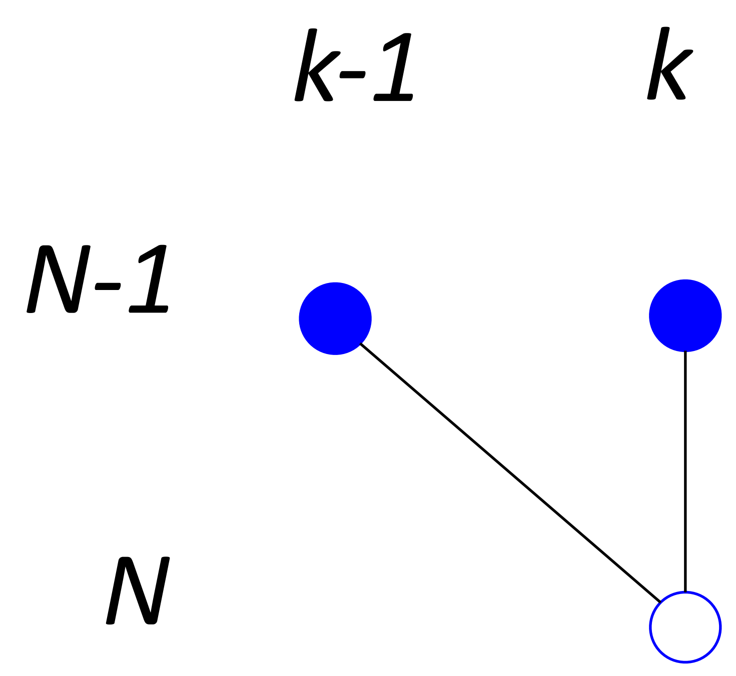
		k = 1	2	3	4	5	6	..
N	1	1	2	2	2	2	2	..
	2	1						
	3	1						
	4	1						
	5	1						
	6	1						
	$\vdots$	$\vdots$						

1. Boundary conditions: easy
2. Induction step

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

The induction step

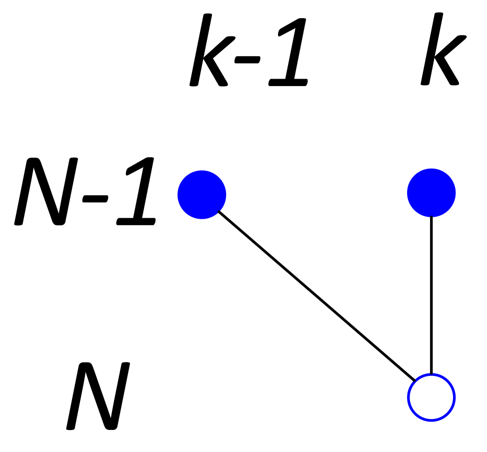
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



$$\boxed{B(N,k) \leq \sum_{i=0}^{k-1} \binom{N}{i}}$$

The induction step

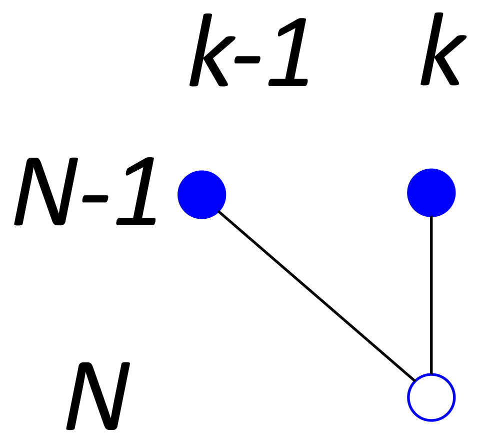
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



$$\boxed{B(N,k) \leq \sum_{i=0}^{k-1} \binom{N}{i}}$$

The induction step

$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$

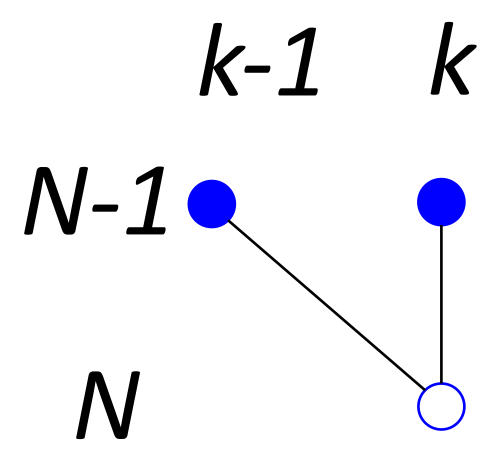


$$\sum_{i=0}^{k-1} \binom{N-1}{i}$$

$$\boxed{B(N,k) \leq \sum_{i=0}^{k-1} \binom{N}{i}}$$

The induction step

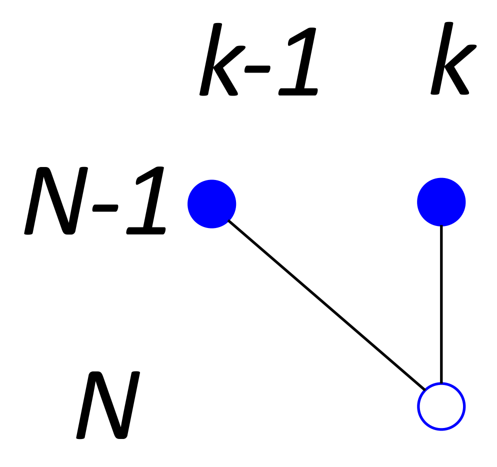
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



$$\sum_{i=0}^{k-1} \binom{N-1}{i} + \sum_{i=0}^{k-2} \binom{N-1}{i}$$

$$\boxed{B(N,k) \leq \sum_{i=0}^{k-1} \binom{N}{i}}$$

The induction step



$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$

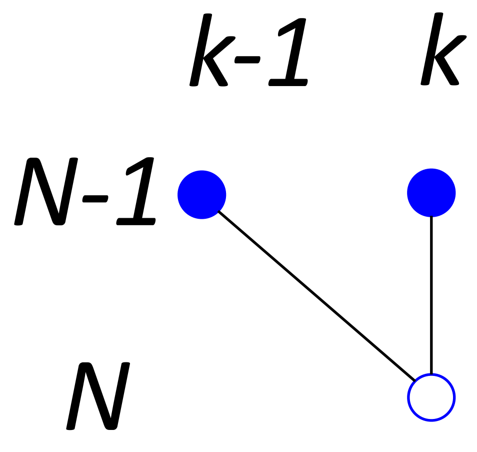
$$\sum_{i=0}^{k-1} \binom{N-1}{i} + \sum_{i=0}^{k-2} \binom{N-1}{i}$$

$$= \quad 1 + \sum_{i=1}^{k-1} \binom{N-1}{i}$$

$$\boxed{B(N,k) \le \sum_{i=0}^{k-1} \binom{N}{i}}$$

The induction step

$$B(N,k) \le B(N-1,k) + B(N-1,k-1)$$



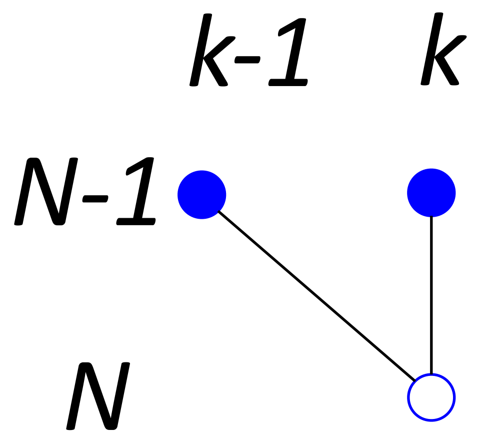
$$\sum_{i=0}^{k-1} \binom{N-1}{i} + \sum_{i=0}^{k-2} \binom{N-1}{i}$$

$$= \quad 1 + \sum_{i=1}^{k-1} \binom{N-1}{i} + \sum_{i=1}^{k-1} \binom{N-1}{i-1}$$

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

The induction step

$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



$$\sum_{i=0}^{k-1} \binom{N-1}{i} + \sum_{i=0}^{k-2} \binom{N-1}{i}$$

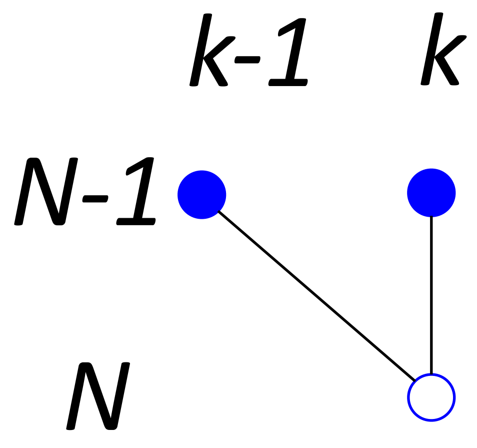
$$= \quad 1 + \sum_{i=1}^{k-1} \binom{N-1}{i} + \sum_{i=1}^{k-1} \binom{N-1}{i-1}$$

$$= \quad 1 + \sum_{i=1}^{k-1} \left[\binom{N-1}{i} + \binom{N-1}{i-1} \right]$$

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

The induction step

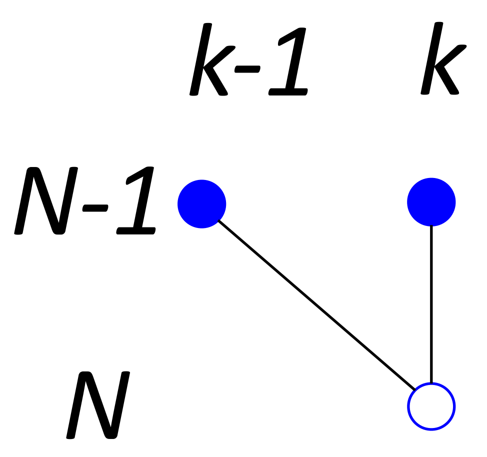
$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$



$$\begin{aligned}
& \sum_{i=0}^{k-1} \binom{N-1}{i} + \sum_{i=0}^{k-2} \binom{N-1}{i} \\
= \quad & 1 + \sum_{i=1}^{k-1} \binom{N-1}{i} + \sum_{i=1}^{k-1} \binom{N-1}{i-1} \\
= \quad & 1 + \sum_{i=1}^{k-1} \left[\binom{N-1}{i} + \binom{N-1}{i-1} \right] \\
= \quad & 1 + \sum_{i=1}^{k-1} \binom{N}{i}
\end{aligned}$$

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

The induction step



$$B(N,k) \leq B(N-1,k) + B(N-1,k-1)$$

$$\begin{aligned}
& \sum_{i=0}^{k-1} \binom{N-1}{i} + \sum_{i=0}^{k-2} \binom{N-1}{i} \\
= \quad & 1 + \sum_{i=1}^{k-1} \binom{N-1}{i} + \sum_{i=1}^{k-1} \binom{N-1}{i-1} \\
= \quad & 1 + \sum_{i=1}^{k-1} \left[\binom{N-1}{i} + \binom{N-1}{i-1} \right] \\
= \quad & 1 + \sum_{i=1}^{k-1} \binom{N}{i} = \sum_{i=0}^{k-1} \binom{N}{i} \quad \checkmark
\end{aligned}$$

It is polynomial!

For a given $\mathcal{H}$, the break point k is fixed

$$m_{\mathcal{H}}(N) \quad \leq$$

It is polynomial!

For a given $\mathcal{H}$, the break point k is fixed

$$m_{\mathcal{H}}(N) \;\leq\; B(N,k) \;\leq\; \sum_{i=0}^{k-1} \binom{N}{i}$$

It is polynomial!

For a given $\mathcal{H}$, the break point k is fixed

$$\boxed{m_{\mathcal{H}}(N) \leq \sum_{i=0}^{k-1} \binom{N}{i}}$$

It is polynomial!

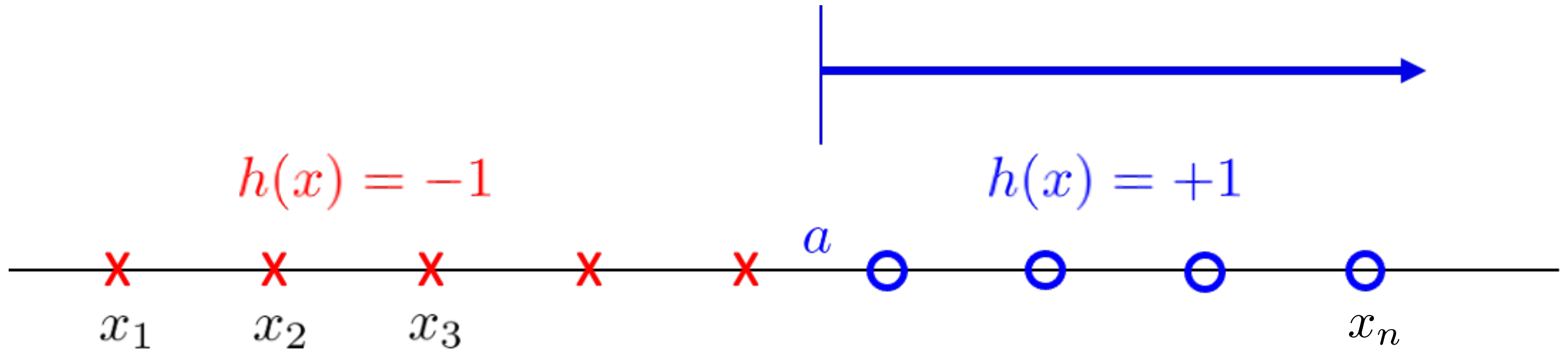
For a given $\mathcal{H}$, the break point k is fixed

$$m_{\mathcal{H}}(N) \leq \underbrace{\sum_{i=0}^{k-1} \binom{N}{i}}_{\text{maximum power is } N^{k-1}}$$

$$\sum_{i=0}^{k-1} \binom{N}{i}$$

3 Examples

- $\mathcal{H}$ is **positive rays** (break point $k = 2$)

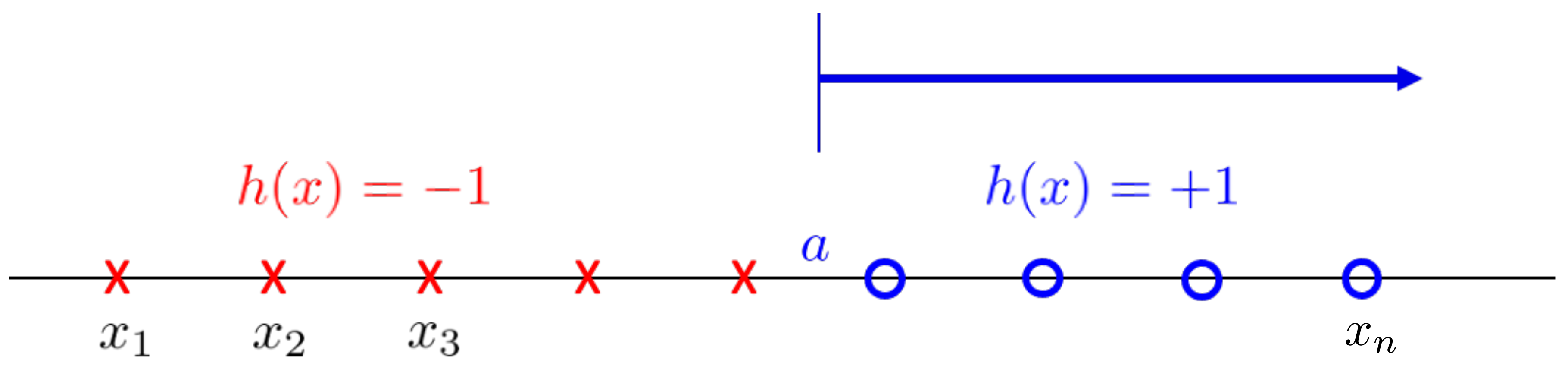


$$m_{\mathcal{H}}(N) \quad = \quad N + 1 \leq N + 1$$

$$\sum_{i=0}^{k-1} \binom{N}{i}$$

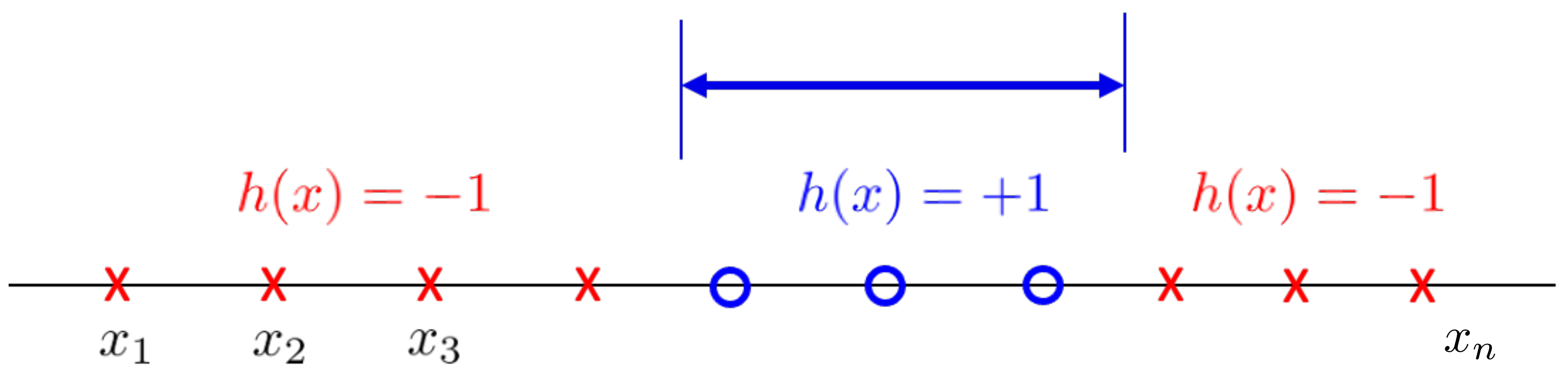
3 Examples

- $\mathcal{H}$ is **positive rays** (break point $k = 2$)



$$m_{\mathcal{H}}(N) \quad = \quad N + 1 \leq N + 1$$

- $\mathcal{H}$ is **positive interval** (break point $k = 3$)

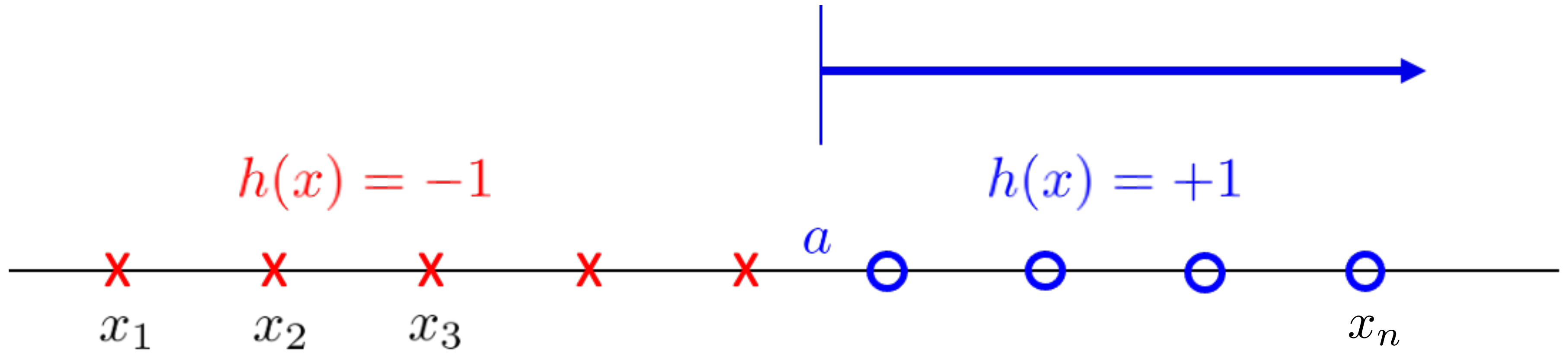


$$m_{\mathcal{H}}(N) \quad = \quad \frac{1}{2}N^2 + \frac{1}{2}N + 1 \leq \frac{1}{2}N^2 + \frac{1}{2}N + 1$$

$$\sum_{i=0}^{k-1} \binom{N}{i}$$

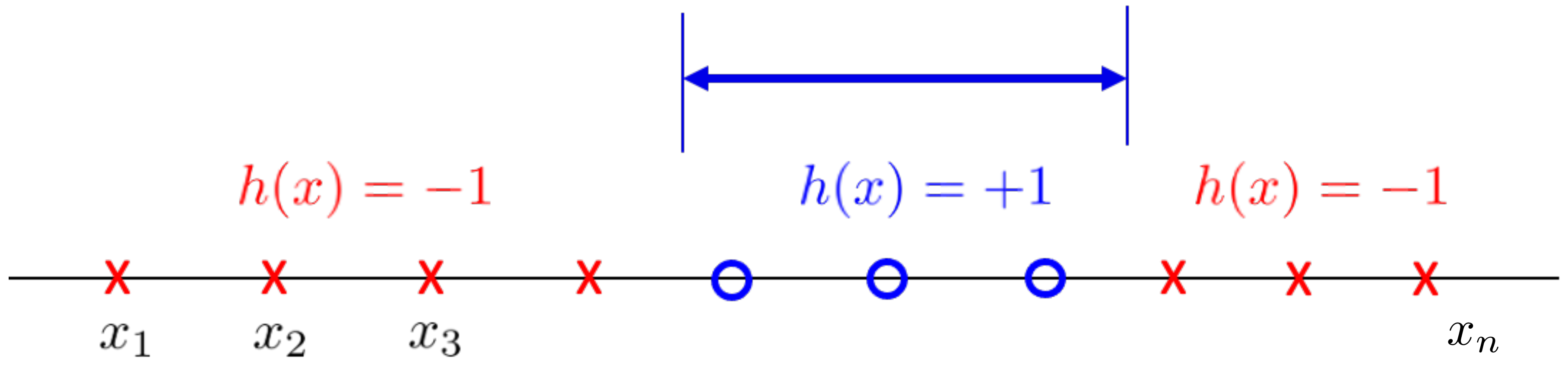
3 Examples

- $\mathcal{H}$ is **positive rays** (break point $k = 2$)



$$m_{\mathcal{H}}(N) \quad = \quad N + 1 \leq N + 1$$

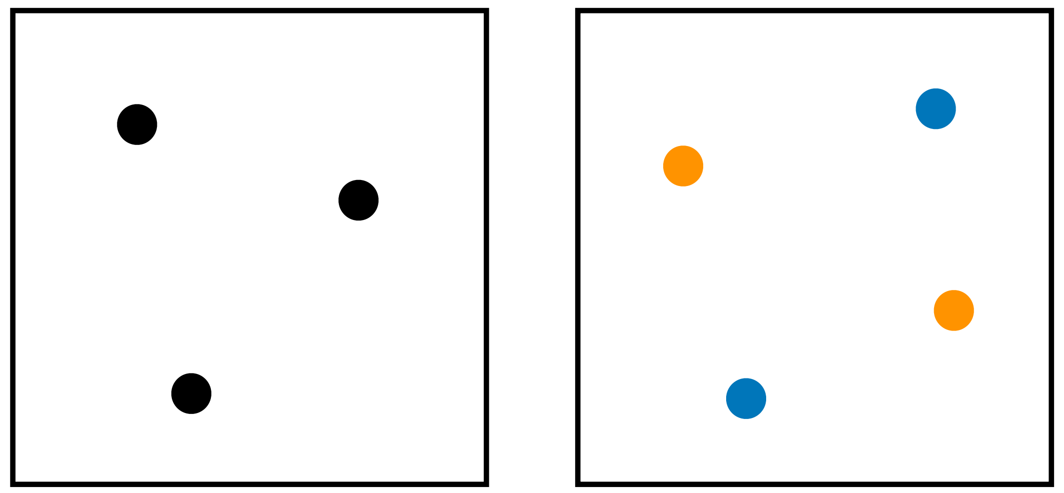
- $\mathcal{H}$ is **positive interval** (break point $k = 3$)



$$m_{\mathcal{H}}(N) \quad = \quad \tfrac{1}{2}N^2 + \tfrac{1}{2}N + 1 \leq \tfrac{1}{2}N^2 + \tfrac{1}{2}N + 1$$

- $\mathcal{H}$ is **2D perceptrons** (break point $k = 4$)

$$m_{\mathcal{H}}(N) \quad = \ ? \leq \tfrac{1}{6}N^3 + \tfrac{5}{6}N + 1$$



$m_{\mathcal{H}}(3) = 8$ $\quad$ $m_{\mathcal{H}}(4) = 14$

Recall: outline

- Prove that $m_{\mathcal{H}}(N)$ is polynomial

- Prove that $m_{\mathcal{H}}(N)$ can replace M

A (tighter) bound with the growth function

Instead of

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \leq 2 \quad M \quad e^{-2N\epsilon^2}$$

we want

$$\mathbb{P}(|\hat{R}_N(h^*) - R(h^*)| > \epsilon) \overset{?}{\leq} 2 \quad m_{\mathcal{H}}(N) \quad e^{-2N\epsilon^2}$$

With probability at least $1 - \delta$: $\quad R(h^*) \overset{?}{\leq} \hat{R}_N(h^*) + \sqrt{\frac{1}{2N} \ln \frac{2m_{\mathcal{H}}(N)}{\delta}}$

Logistics

- Quiz on Thursday
- Based on HW1 and HW2
- 15 mins in-class
- Closed book