VC Dimension
Midterm Recap

PAC Bounds

- Suppose a consistent learner has hypothesis space *H*.
- Then for any *h* in *H* with 0 training error,

$$P\left(error_{true}(h) \leq m^{-1}\left[\ln|H| - \ln(\delta)\right]\right) > \delta$$

- *m* is # training examples,
- |*H*| measures the 'flexibility' of the classifier.
- Problem: Assumes discrete *H*

VC Dimension

- VC(*H*) also measures the 'flexibility' of *H*.

$$P(\epsilon \geq m^{-1}(4\log_2(2/\delta)+8\mathrm{VC}(H)\log_2(13/\epsilon))) < \delta$$

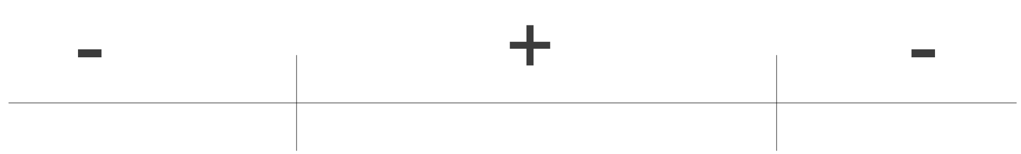
- *m* is # training examples
- ϵ is desired true error
- δ is probability of achieving desired true error

Computing VC Dimension

- Size of largest set set of instances that can be 'shattered' by a classifier
- A classifier can 'shatter' a set of *n* instances if the classifier can assign every possible labeling in $\{0,1\}^n$ to those examples.
- 'instance' is the X part of a training or testing example

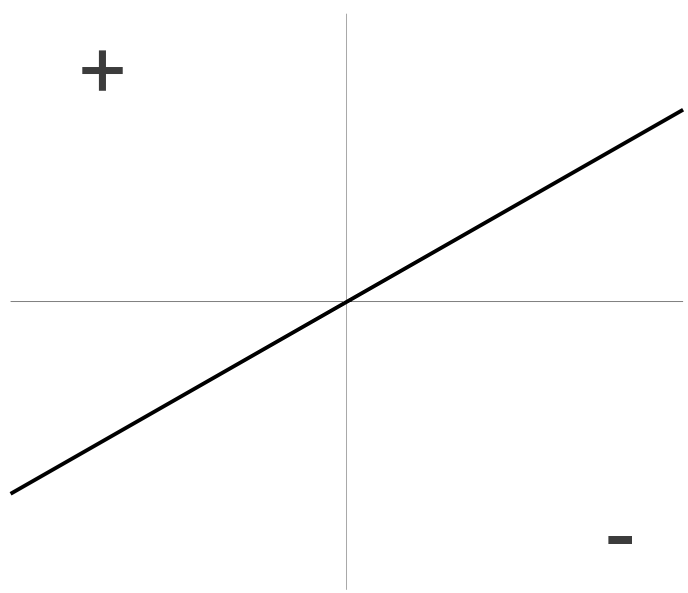
Example I

- Instances: $\in \mathbb{R}^1$
- Hypothesis space:
  - parameterized by {*a*,*b*}
  - classify *x* as 1 iff $a<x<b$



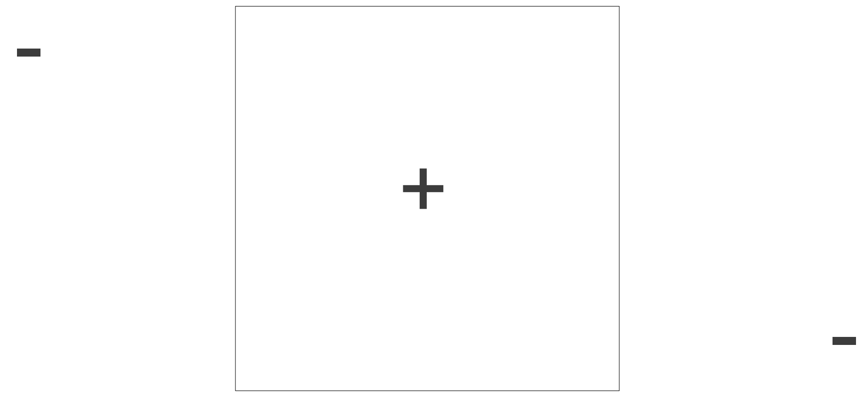
Example II

- Instances : $\in \mathbb{R}^2$
- Hypothesis space:
  - parameterized by {*a*,*b*}
  - classify (x,y) as 1 iff *ax*+*by*>0



Example III

- Instances: $\in \mathbb{R}^2$
- Hypothesis space:
  - parameterized by {*a*,*b*,*s*}
  - classify *x* as 1 iff *x* is inside the the square of side *s* centered at (*a*,*b*).



Example IV

- Instances: $\in \{0,1\}^{10}$
- Hypothesis space:
  - Decision Trees on *n* variables

Midterm: Tough Questions

- 1.1.1-2: Does NB achieve 0 train/test error?
- 1.1.4-5: Are two Bayes nets equivalent?
- 2.1.6: EM Algorithm on a Bayes net
- 2.2: Constructing a Bayes net
- 4: Linear regression
- 5.1-2: Naive Bayes with cond. indep. violation
- 6: Extra Credit: Violated Bayes net assumption