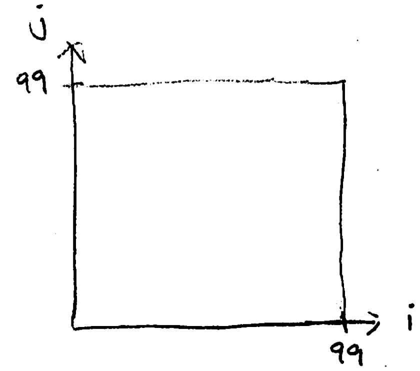
Byron Becker

Problem 1



$100 \times \underline{100}$ space $= 10^4$ options

so $\binom{10^4}{3}$ possibilities b/c can choose any 2 points given the first choice (collinear is allowed)

so $|H| = \binom{10^4}{3} = \frac{10^4!}{(10^4-3)!\,3!} = 1.66617 \times 10^{11}$

lower bound on training examples for a finite hypothesis class

$$m \geq \frac{1}{\varepsilon}\left(\ln|H| + \ln\left(\frac{1}{\delta}\right)\right)$$

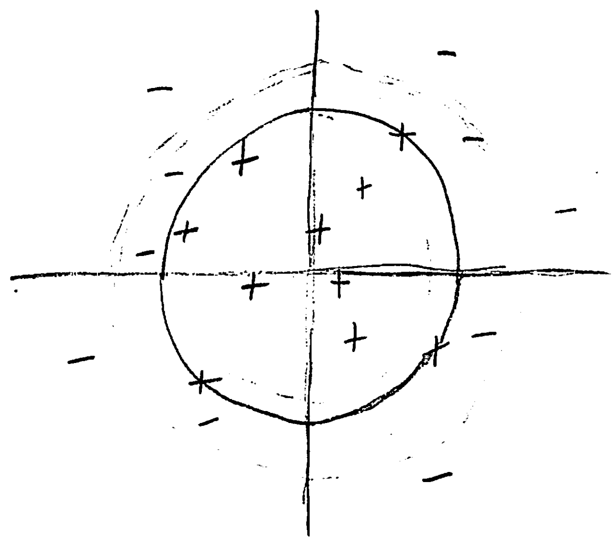
want to output an error of at most 15% w/95% confidence

so $\varepsilon = 0.15$, $\delta = 1 - 0.95 = 0.05$

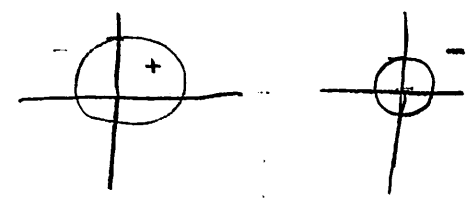
$$m \geq \frac{1}{0.15}\left(\ln\left(1.67 \times 10^{11}\right) + \ln\left(\frac{1}{0.05}\right)\right) = 192.2 \simeq 192 \text{ training examples}$$

Problem 2

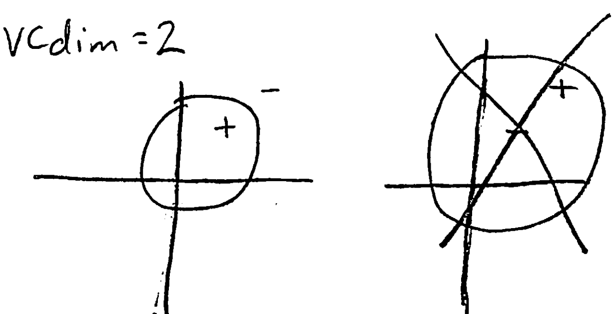
Circles centered at the origin. A hypothesis h in this class can either classify points as positive if they lie on the boundary or interior of the circle, or can classify points as positive if they lie on the boundary or exterior of the circle. State & prove the VC dimension of this family of classifiers



if VCdim = 1

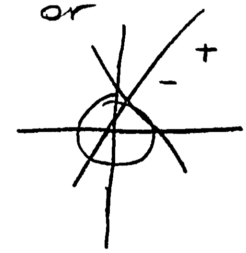


if VCdim = 2



so graphically we see the VCdim = 1

or



Proof

consider 1 pt p_1 w/ $r_1 > 0$ and $0 \le \theta_1 < 2\pi$

if $p_1 = \oplus$, then draw a circle w/ $r_c \ge r_1$ centered at 0 w/ eqn $P(r_c, \theta) = r_c \cos\theta + r_c \sin\theta$ therefore, p_1 will fall w/in this circle or on its boundary.

if $p_1 = \ominus$, then draw a similar circle w/ $r_c < r_1$, therefore, p_1 will lie outside the circle's boundary

Now consider 2pts p_1, p_2 w/ $r_1 \le r_2$

if $p_1 = \ominus$ and $p_2 = \oplus$, then

if $r_c < r_1$, we have a circle that correctly excludes p_1 but incorrectly excludes p_2

visa versa, if $r_c > r_2$, we have a circle that correctly includes p_2, but as $r_2 > r_1$, incorrectly includes p_1. Therefore we cannot shatter a VCdim of 2, so the VCdim = 1

Problem 3

My Intuition
If we have data points j = 0 to n, then for their positions i, all data points

$$x_i = 2^{-i}$$

gives an infinite possible number of points between 0 and 1. For our hypothesis $h_w(x)$, looking at the function $y = \sin(wx)$, w relates to the period of our sine wave by $p = \frac{2\pi}{w}$. This means that $y = \sin(\pi x)$ has a period of 2.

Looking further at the function, we can see that for multiples of 2, our function will pass through the points $x = 2^{-i}$ at $y = 0$. This means that if we are to correctly classify points with

$$if \quad y = sin(wx) < 0 \quad then \ -1 \ or \ False$$
$$else \quad 1 \ or \ True$$

then we have to implement a shift in the period with a small enough magnitude to not incorrectly classify our smallest points. To do this, I looked at the graph of $y = \sin(2\pi x)$ and then implemented an initial shift to the period of 0.01 by changing the equation to $y = \sin((2\pi + 0.01)x)$. Looking in excel at different points $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, and $\frac{1}{16}$ classified by different functions based on $y = \sin(wx)$ with

$$w = (\sum_{i=1}^{n} 2^i) + 2^n$$

I found the following pattern.

	2^3	2^3+2^1	2^3+2^2	2^3+2^2+2^1	2^4
	8	10	12	14	16
0.5	TRUE	FALSE	TRUE	FALSE	TRUE
0.25	TRUE	TRUE	FALSE	TRUE	TRUE
0.125	FALSE	FALSE	FALSE	TRUE	TRUE

	2^4	2^4+2^1	2^4+2^2	2^4+2^2+2^1	2^4+2^3	2^4+2^3+2^1	2^4+2^3+2^2	2^4+2^3+2^2+2^1	2^5
	16	18	20	22	24	26	28	30	32
0.5	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE
0.25	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE
0.125	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE
0.0625	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE

	2^5	2^5+2^1	2^5+2^2	2^5+2^2+2^1	2^5+2^3	2^5+2^3+2^1	2^5+2^3+2^2	2^5+2^3+2^2+2^1	2^5+2^4	etc...							
	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62	64
0.5	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE
0.25	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE
0.125	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE
0.0625	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE
0.03125	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE

Therefore, we can see a somewhat binary pattern emerging, almost one like the one in bits. As we increase by 2, we see higher "level" i bits turn to False (meaning the function evaluates to less than 0 at that point, just like in binary). Looking at this pattern one can see that for each false or -1 result, we have a

$$\sum_{i=1}^{n} 2^i$$

pattern emerging where the i's are the subset of numbers associated in tuples with the false (or -1) classification.

To then make sure that I was shifting the period of the wave by a correct amount in all cases I followed the sum pattern I found in the excel sheet jpg, and made sure to shift the period of the sin wave by the maximum i (or minimum i depending on how you look at it) as follows

$$w = (\sum_{i=1}^{n} 2^i) + 2^{-n}$$

where again the i's are the subset of numbers associated in tuples with the false (or -1) classification. I found that both from my testing using wolfram alpha and in the provided tests that this 2^{-n} provided a sufficient shift to the left as the shift of the period as it would shift from including that point in the positive class to the negative class.

4 Distinct Points that Cannot Be Shattered?

Since a sine wave has a consistent period, we can assume that for any x, $f(x) == f(x + multiple * period)$ of the sine wave. Therefore, given the points $x = \frac{1}{2}, 1, \frac{3}{2}$, and 2, if we classify $\frac{1}{2}$, 1, and 2 as false but $\frac{3}{2}$ as true, if we set the period to $w = 2\pi + 2^{-1}$ to correctly classify the point at $x = \frac{1}{2}$, we will incorrectly classify the point at $x = \frac{3}{2}$ because they are separated exactly by some period multiple of the sin wave.

Plot: