DSC 190/291 · Assignment 2 Solutions

UCSD · Spring 2026

Part A: Unions of Two Intervals on the Line (40 points)

Let

$$\mathcal{H}_2 = \{x \to \mathbf{1}[x \in I_1 \cup I_2] : I_1, I_2 \subseteq \mathbb{R} \text{ are intervals}\}.$$

In other words, a hypothesis in $\mathcal{H}_2$ labels a point by 1 iff it lies in a union of at most two intervals on the real line. Empty intervals are allowed, so this really means "at most two."

Throughout this part, fix ordered points

$$x_1 < x_2 < \ldots < x_n.$$

1. **Restriction patterns.**

   Characterize exactly which binary labelings of $(x_1, \ldots, x_n)$ can be realized by $\mathcal{H}_2$.

   A good answer should not just list examples. It should identify the structural property that distinguishes realizable labelings from impossible ones.

   Solution.

   A labeling $(y_1, \ldots, y_n) \in \{0,1\}^n$ is realizable by $\mathcal{H}_2$ if and only if the set of indices with label 1 is the union of at most two contiguous blocks.

   To see necessity, fix $h \in \mathcal{H}_2$. Each interval I intersects the ordered pool in a contiguous block of indices: if $x_i, x_k \in I$, then for any index j with $i < j < k$, we also have $x_j \in I$ because I is an interval. Therefore $I_1 \cup I_2$ can produce at most two contiguous blocks of positive labels.

   Conversely, suppose the 1 labels form at most two contiguous blocks. If there are no positive labels, choose both intervals empty. If there is one block $x_i, x_{i+1}, \ldots, x_j$, choose $I_1 = [x_i, x_j]$ and I_2 empty. If there are two blocks $x_i, \ldots, x_j$ and $x_k, \ldots, x_l$ with $j < k$, choose $I_1 = [x_i, x_j]$ and $I_2 = [x_k, x_l]$. These intervals realize exactly the desired labels on the fixed ordered pool.

2. **Exact growth function.**

   Use your characterization to derive a closed-form formula for the growth function

   $$\Gamma_{\mathcal{H}_2}(n).$$

   You may first use small cases to guess the formula, but your final answer must be a proof.

   Solution.

By the characterization above, we count binary strings of length n whose 1 labels form zero, one, or two contiguous blocks.

There is exactly 1 pattern with zero positive blocks: the all-zero pattern.

A one-block pattern is determined by choosing the first and last positive positions. Equivalently, choose two cut positions among the $n+1$ gaps around the ordered points. This gives

$$\binom{n+1}{2}$$

patterns.

A two-block pattern has the form

$$0...01...10...01...10...0,$$

with both 1 blocks nonempty and with a nonempty zero block between them. It is determined by four cut positions

$$0 \le p < q < r < s \le n,$$

where the positive blocks are $x_{p+1}, ..., x_q$ and $x_{r+1}, ..., x_s$. Thus there are

$$\binom{n+1}{4}$$

two-block patterns.

Therefore

$$\Gamma_{\mathcal{H}_2}(n) = 1 + \binom{n+1}{2} + \binom{n+1}{4}.$$

Using Pascal's identity,

$$\binom{n+1}{2} = \binom{n}{1} + \binom{n}{2}$$

and

$$\binom{n+1}{4} = \binom{n}{3} + \binom{n}{4},$$

so equivalently

$$\Gamma_{\mathcal{H}_2}(n) = \sum_{k=0}^{4} \binom{n}{k}.$$

3. **VC dimension and mistake bound.**

Determine the exact value of

$$\mathrm{VCdim}(\mathcal{H}_2).$$

Then use your exact growth formula to write down the Halving mistake bound on an n-point pool in the realizable online transductive setting.

Finally, compare your exact growth formula with the Sauer–Shelah bound obtained from your VC-dimension calculation. Are they equal in this case? What does this tell you about when a VC-dimension-based upper bound can be tight?

Solution.

The exact VC dimension is

$$\mathrm{VCdim}(\mathcal{H}_2) = 4.$$

For any four ordered points, the growth formula gives

$$\Gamma_{\mathcal{H}_2}(4) = \sum_{k=0}^{4} \binom{4}{k} = 2^4,$$

so four points are shattered.

For any five ordered points,

$$\Gamma_{\mathcal{H}_2}(5) = \sum_{k=0}^{4} \binom{5}{k} = 31 < 32 = 2^5.$$

Thus no five-point pool can be shattered. Hence the VC dimension is exactly 4.

In the realizable online transductive setting on an n-point pool, Halving on restrictions makes at most

$$\log_2 \Gamma_{\mathcal{H}_2}(n) = \log_2 \left(\sum_{k=0}^{4} \binom{n}{k} \right)$$

mistakes.

Since $\mathrm{VCdim}(\mathcal{H}_2) = 4$, Sauer–Shelah gives

$$\Gamma_{\mathcal{H}_2}(n) \le \sum_{k=0}^{4} \binom{n}{k}.$$

In this case the inequality is an equality for every n. Thus the VC-dimension-based upper bound can be exactly tight, not merely correct up to order, for some classes.

4. **Tightness of Sauer–Shelah.**

For arbitrary integers $n \ge d \ge 0$, give a concrete hypothesis class $\mathcal{H}$ on some domain $\mathcal{X}$ such that

$$\mathrm{VCdim}(\mathcal{H}) = d$$

and

$$\Gamma_{\mathcal{H}}(n) = \sum_{k=0}^{d} \binom{n}{k}.$$

What does your example show about the Sauer–Shelah bound?

Solution.

Let

$$\mathcal{X} = \{1, 2, ..., n\}$$

and let $\mathcal{H}$ be the class of indicators of subsets of size at most d:

$$\mathcal{H} = \{x \to \mathbf{1}[x \in A] : A \subseteq \mathcal{X}, |A| \leq d\}.$$

First, $\text{VCdim}(\mathcal{H}) \geq d$ because any d points are shattered: every labeling of those d points corresponds to choosing the positively labeled subset, whose size is at most d.

Second, $\text{VCdim}(\mathcal{H}) \leq d$. No set of $d+1$ points can be shattered, because the labeling that assigns 1 to all $d+1$ points would require a set A of size at least $d+1$, which is not allowed.

Therefore $\text{VCdim}(\mathcal{H}) = d$.

On the full n-point domain, the distinct patterns are exactly the subsets of $\mathcal{X}$ of size at most d, so

$$\Gamma_{\mathcal{H}}(n) = \sum_{k=0}^{d} \binom{n}{k}.$$

This shows that the Sauer–Shelah bound is best possible in general. There are classes for which the upper bound is achieved exactly.

5. **AI proof audit.**

An AI assistant claims:

> On an ordered n-point sample, a labeling realized by $\mathcal{H}_2$ is determined by the places where the labels switch between 0 and 1. Since a union of at most two intervals can create at most four such switches, one just chooses up to four switch locations among the n sample positions. Therefore
>
> $$\sum_{j=0}^{4} \binom{n}{j}.$$
>
> Hence
>
> $$\Gamma_{\mathcal{H}_2}(n) = \sum_{j=0}^{4} \binom{n}{j},$$
>
> and in particular
>
> $$\text{VCdim}(\mathcal{H}_2) = 4.$$

A flawed argument may still arrive at a true conclusion; analyze the reasoning, not just the final claim.

Explain carefully what is incomplete or incorrect in this argument. Then replace it with a correct statement that is actually supported by your work above.

Solution.

The final formula is true, but the argument as written does not justify it.

First, the phrase "switch locations among the n sample positions" is not precise. Adjacent label switches occur in the $n-1$ gaps between sample points, while block endpoints are more naturally counted using the $n+1$ cut positions around the sample. These are different parameterizations, and the argument does not specify which one is being used.

Second, "at most four switches" alone is not a complete structural characterization. For example, the string 10101 has four adjacent switches, but it has three positive blocks and cannot be realized by a union of two intervals. What matters is that the 1 labels form at most two contiguous blocks.

Third, even after finding an upper bound, one still needs to prove that every counted pattern is actually realizable by two intervals. That converse direction is missing.

A correct statement is: on any ordered n-point pool, $\mathcal{H}_2$ realizes exactly the binary strings whose 1 labels form at most two contiguous blocks. Therefore

$$\Gamma_{\mathcal{H}_2}(n) = 1 + \binom{n+1}{2} + \binom{n+1}{4} = \sum_{k=0}^{4} \binom{n}{k},$$

and hence

$$\mathrm{VCdim}(\mathcal{H}_2) = 4.$$

The proof is the block characterization and counting argument from the previous parts.

Part B: Quadratic Threshold Functions on the Line (45 points)

Let

$$\mathcal{H}_{\mathrm{quad}} = \{x \to \mathbf{1}[ax^2 + bx + c \geq 0] : (a, b, c) \in \mathbb{R}^3, (a, b, c) \neq (0, 0, 0)\}.$$

So a hypothesis in $\mathcal{H}_{\mathrm{quad}}$ is the indicator of the nonnegative region of a quadratic polynomial.

Throughout this part, again fix ordered points

$$x_1 < x_2 < \ldots < x_n.$$

1. **A general upper-bound trick.**

   Prove the following statement.

   If there exist an integer $D \geq 1$ and a transformation

   $$\varphi : \mathcal{X} \to \mathbb{R}^D$$

   such that every hypothesis $h \in \mathcal{H}$ can be written in the form

   $$h(x) = \mathbf{1}[\langle w, \varphi(x) \rangle \geq 0]$$

for some vector $w \in \mathbb{R}^D$, then

$$\mathrm{VCdim}(\mathcal{H}) \leq D.$$

You may use the Week 2 result that homogeneous halfspaces in $\mathbb{R}^D$ have VC dimension D.

Solution.

Let $\mathcal{G}$ be the class of homogeneous halfspaces in $\mathbb{R}^D$:

$$\mathcal{G} = \{z \to \mathbf{1}[\langle w, z \rangle \geq 0] : w \in \mathbb{R}^D\}.$$

Suppose, toward a contradiction, that $\mathcal{H}$ shatters a set $C = \{x_1, ..., x_m\} \subseteq \mathcal{X}$ with $m > D$. Then for every labeling $(y_1, ..., y_m) \in \{0, 1\}^m$, there is some $h \in \mathcal{H}$ such that $h(x_i) = y_i$ for all i.

By assumption, this h can be written as $h(x) = \mathbf{1}[\langle w, \varphi(x) \rangle \geq 0]$ for some $w \in \mathbb{R}^D$. Therefore the homogeneous halfspace with normal vector w realizes the labeling $(y_1, ..., y_m)$ on the transformed points $\varphi(x_1), ..., \varphi(x_m)$.

The transformed points must be distinct: if $\varphi(x_i) = \varphi(x_j)$ for $i \neq j$, then no homogeneous halfspace could assign different labels to x_i and x_j, contradicting shattering of C by $\mathcal{H}$. Thus homogeneous halfspaces in $\mathbb{R}^D$ would shatter m transformed points. This contradicts the Week 2 result that homogeneous halfspaces in $\mathbb{R}^D$ have VC dimension D. Hence no set of size larger than D can be shattered by $\mathcal{H}$, so

$$\mathrm{VCdim}(\mathcal{H}) \leq D.$$

2. **Apply the trick to quadratic thresholds.**

   Find an explicit transformation

$$\varphi : \mathbb{R} \to \mathbb{R}^3$$

   such that every hypothesis in $\mathcal{H}_{\text{quad}}$ can be written in the form

$$h(x) = \mathbf{1}[\langle w, \varphi(x) \rangle \geq 0]$$

   for some $w \in \mathbb{R}^3$.

   Then use the previous item to prove an upper bound on

$$\mathrm{VCdim}(\mathcal{H}_{\text{quad}}).$$

Solution.

Use

$$\varphi(x) = (x^2, x, 1) \in \mathbb{R}^3.$$

For $w = (a, b, c) \in \mathbb{R}^3$, we have

$$\langle w, \varphi(x) \rangle = ax^2 + bx + c.$$

Therefore every quadratic threshold function can be written as

$$h(x) = \mathbf{1}[\langle w, \varphi(x) \rangle \geq 0].$$

Applying the previous part with $D = 3$ gives

$$\text{VCdim}(\mathcal{H}_{\text{quad}}) \leq 3.$$

3. **Exact restriction patterns.**

Characterize exactly which binary labelings of $(x_1, ..., x_n)$ are realizable by $\mathcal{H}_{\text{quad}}$.

Your characterization should make clear how the algebraic structure of a quadratic polynomial controls the combinatorics of its restriction to an ordered finite set.

Solution.

A labeling $(y_1, ..., y_n) \in \{0, 1\}^n$ is realizable by $\mathcal{H}_{\text{quad}}$ if and only if it has at most two adjacent label changes. Equivalently, either the 1 labels form a single contiguous block, or the 0 labels form a single contiguous block, allowing the block to be empty or to touch the boundary.

For necessity, let $q(x) = ax^2 + bx + c$ be a nonzero polynomial of degree at most 2. The set $\{x : q(x) \geq 0\}$ is empty, all of $\mathbb{R}$, a ray, a closed interval, the complement of an open interval, or a single point. In each case, its restriction to an ordered finite sample has at most two adjacent changes. Equivalently, changes can occur only at real roots of q, counted with multiplicity, and such a polynomial has at most two of them.

For sufficiency, take any binary string with at most two adjacent changes.

If there are no changes, use the constant polynomial $q(x) = 1$ for the all-one string and $q(x) = -1$ for the all-zero string.

If there is one change, the string is a threshold pattern. If it has the form 0...01...1, choose θ in the gap where the change occurs and use $q(x) = x - \theta$. If it has the form 1...10...0, use $q(x) = \theta - x$.

If there are two changes, then the string has one of the forms 0...01...10...0 or 1...10...01...1. In the first case, choose numbers $\alpha < \beta$ in the two gaps where the labels change and use $q(x) = -(x - \alpha)(x - \beta)$, which is nonnegative between α and β. In the second case, use $q(x) = (x - \alpha)(x - \beta)$, which is nonnegative outside the interval between α and β.

Therefore the exact restriction patterns are precisely the binary strings with at most two adjacent changes.

4. **Exact growth and exact VC dimension.**

Use your characterization to compute the exact growth function

$$\Gamma_{\mathcal{H}_{\text{quad}}}(n)$$

and the exact VC dimension of $\mathcal{H}_{\text{quad}}$.

Then compare your exact growth formula with the Sauer–Shelah bound obtained from your VC-dimension calculation. Is the VC-based upper bound tight here? Explain what your answer says about VC dimension as a summary of finite-pool richness.

Solution.

For $n \geq 1$, a binary string with at most two adjacent changes is determined by:

- the starting label $y_1 \in \{0, 1\}$, and
- the set of adjacent gaps where the label changes, of size 0, 1, or 2.

There are $n - 1$ adjacent gaps, so

$$\Gamma_{\mathcal{H}_{\text{quad}}}(n) = 2 \sum_{j=0}^{2} \binom{n-1}{j}.$$

Equivalently,

$$\Gamma_{\mathcal{H}_{\text{quad}}}(n) = 2\left(1 + (n-1) + \binom{n-1}{2}\right) = n^2 - n + 2.$$

For the empty pool, one may take $\Gamma_{\mathcal{H}_{\text{quad}}}(0) = 1$.

The exact VC dimension is 3. Indeed,

$$\Gamma_{\mathcal{H}_{\text{quad}}}(3) = 2 \sum_{j=0}^{2} \binom{2}{j} = 8 = 2^3,$$

so any three ordered points are shattered. But

$$\Gamma_{\mathcal{H}_{\text{quad}}}(4) = 2 \sum_{j=0}^{2} \binom{3}{j} = 14 < 16 = 2^4,$$

so no four-point pool is shattered.

Thus

$$\mathrm{VCdim}\big(\mathcal{H}_{\text{quad}}\big) = 3.$$

This also matches the upper bound from the feature-map argument.

Sauer–Shelah with VC dimension 3 gives

$$\Gamma_{\mathcal{H}_{\text{quad}}}(n) \leq \sum_{k=0}^{3} \binom{n}{k}.$$

This is not equal to the exact growth function for all n. For example, at $n = 4$ the exact value is 14, while the Sauer–Shelah bound is 15.

Thus the VC dimension gives a general upper bound on finite-pool richness, but it need not determine the exact growth function. Two classes with the same VC dimension can have different growth functions above the VC dimension.

5. **AI proof audit.**

An AI assistant claims:

A quadratic polynomial has at most two real roots, so on an ordered sample its labels can change from 0 to 1 or from 1 to 0 at most twice. Therefore one chooses up to two change-points among the $n-1$ gaps and gets

$$\Gamma_{\mathcal{H}_{\text{quad}}}(n) = \sum_{j=0}^{2} \binom{n-1}{j}.$$

In particular,

$$\mathrm{VCdim}(\mathcal{H}_{\text{quad}}) = 3.$$

A flawed argument may still arrive at a true conclusion; analyze the reasoning, not just the final claim.

Explain carefully what is incomplete or incorrect in this argument. Then replace it with a correct theorem that is genuinely justified by your work in this part.

Solution.

The argument correctly identifies an important constraint: a quadratic polynomial has at most two real roots, so the labels along an ordered sample can change at most twice.

However, the counting is incomplete. Choosing the change-points does not determine the whole labeling unless one also chooses the starting label. For example, with no change-points there are two possible strings, all zeros and all ones. The claimed formula counts only one of them.

Also, the claimed VC-dimension conclusion does not follow from the claimed growth formula. If

$$\Gamma_{\mathcal{H}_{\text{quad}}}(n)$$

were really

$$\sum_{j=0}^{2} \binom{n-1}{j},$$

then at $n = 3$ it would equal 4, not 8, so it would not show that three points are shattered.

The correct theorem is: on every ordered n-point pool with $n \geq 1$, $\mathcal{H}_{\text{quad}}$ realizes exactly the binary strings with at most two adjacent label changes. Therefore

$$\Gamma_{\mathcal{H}_{\text{quad}}}(n) = 2\sum_{j=0}^{2} \binom{n-1}{j} = n^2 - n + 2,$$

and

$$\mathrm{VCdim}(\mathcal{H}_{\text{quad}}) = 3.$$