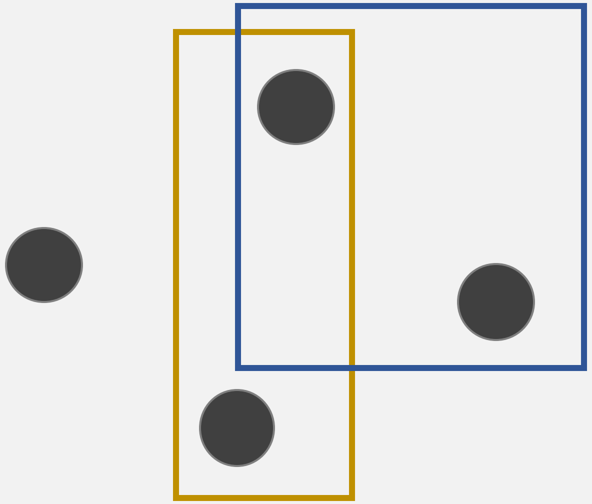
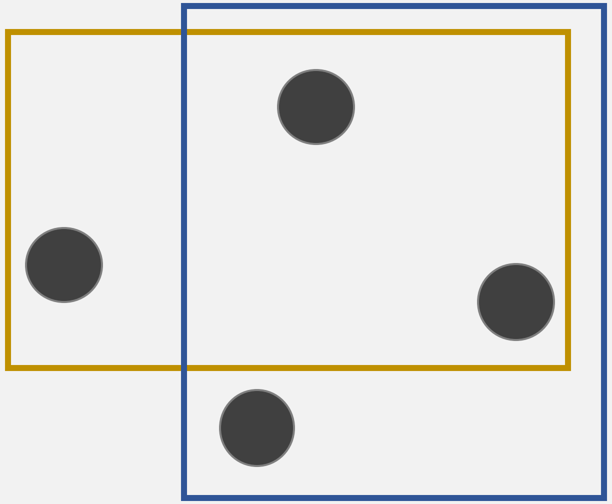
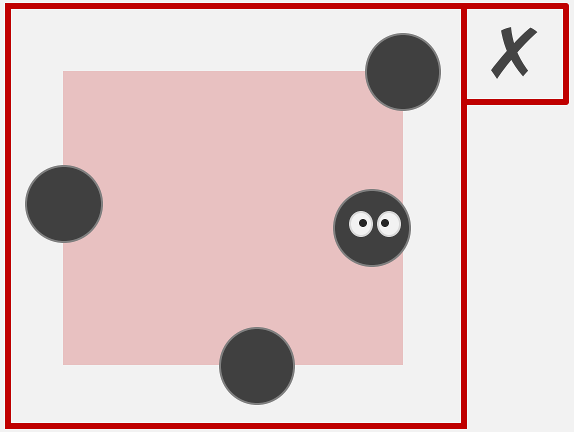
1. Suppose the input dimension is 2. Show that the VC dimension of the hypothesis class of axis-aligned rectangles that include positive instances inside and exclude negative instances outside is 4.

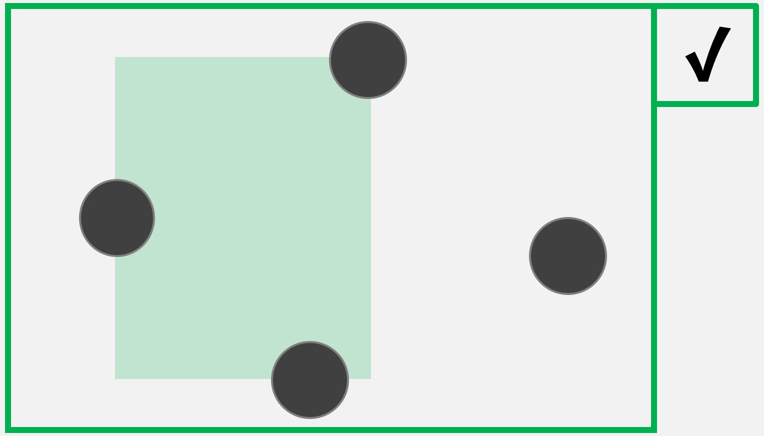
- 在2維平面上，存在一種4個點的排列方式，使得矩形可以包含它們之中的任意1、2、3、4個點
- 如下的4個點的排列方式，即滿足以上的性質
- 首先，我們知道「矩形的VC dimension至少有4」



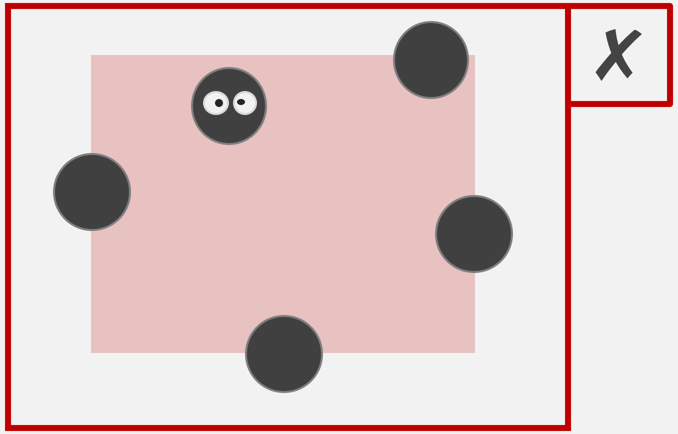


- 若一種4個點的排列方式，使得矩形可以包含它們之中的任意1、2、3、4個點，那麼顯然對它們中的任何一點而言，都不能落在可以包含其他3個點的最小矩形內。



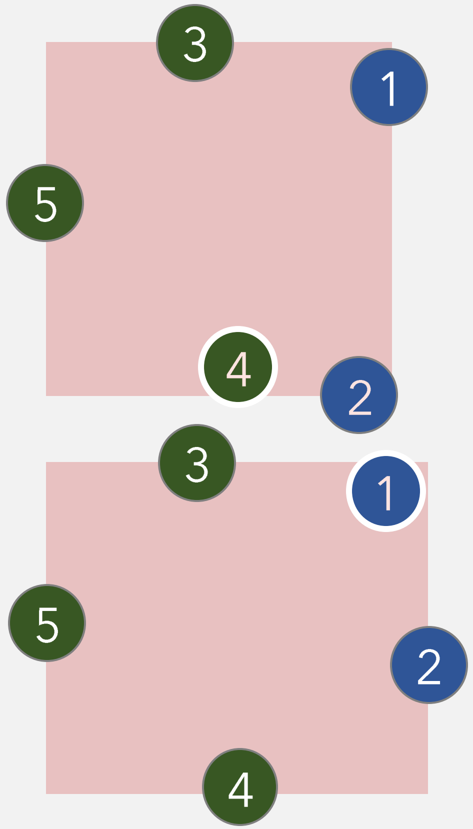


- 同理，我們需要證明任何一種5個點的排列方式，必然可以在其中找到1個點，落在可以包含其他4個點所形成的最小矩形內



- 如前所述，這5的點至少必須形成一個凸5邊形。
- 我們固定③、④、⑤這3個點，事實上，這樣是合理的，我們甚至可以假設這個2維平面的基底是由這三個點定義的。
- 並且假設①、②都在這3個點的右方(X值較這3個點為大)。這也是合理假設，否則不可能以矩形分開最右邊的2個點。
- 因此這些看似特殊的假設實際上是具有普遍性的，如右圖。

- 那麼，以(x_1 , y_1)代表①的位置，(x_2 , y_2)代表②的位置，依此類推
- 如右圖，必須$y_1 < y_3$， $y_2 > y_4$。
- 很明顯地，不可能同時存在以下兩種情況，即：
  1. $x_1 > x_2$
  2. $x_2 > x_1$
- 任意排列的5個點，矩形必然無法排除其中的1個點
- 因此矩形的VC dimension小於5
- 綜上所述，矩形的VC dimension為4



我們換個進階的方式來說明

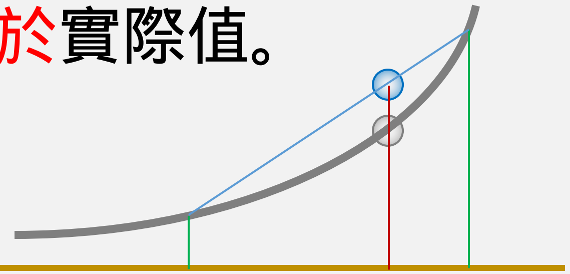
聚焦於可以包含所有點的最小矩形

1. 4個點(A, B, C, D)可以決定一個四邊平行於X軸、Y軸的矩形，包含A, B, C, D，且面積為最小
2. 因此，加入第5個點(E)之後，若E
    1. 落在這個矩形內，則無法以任何矩形僅排除E而包含A, B, C, D，因為包含A, B, C, D的最小矩形也包含E。
    2. 落在這個矩形外，那麼矩形必須擴張來包含E ，擴張的過程必然會使得原先A, B, C, D中的其中一個點落在擴張後的矩形內，假設這個點是A。如此一來，則無法以任何矩形僅排除A而包含B, C, D, E，因為包含B, C, D, E的最小矩形也包含A。

2. Read the proof of Hoeffding's inequality, and show that given a random variable Z and a real-valued function f with values $f(Z) \in [a, b]$, for any real number η, we have

$$E[e^{\eta f(Z)}] \le \frac{b - E[f(Z)]}{b-a} e^{\eta a} + \frac{E[f(Z)] - a}{b-a} e^{\eta b}.$$

- $a \le f(Z) \le b$，因此 $a \le E[f(Z)] \le b$， 給定Z的情況下，$E[f(Z)]$可視為常數
- 對exponential 函數進行線性內插估計，其估計值總是會大於或等於實際值。以下的兩個操作都等效於線性內插估計。



1. $e^{\eta a} = E[e^{\eta a}] \le E[e^{\eta f(Z)}] \le E[e^{\eta b}] = e^{\eta b}$，且 $e^{\eta a} \le e^{\eta E[f(Z)]} \le e^{\eta b}$，因此$e^{\eta E[f(Z)]}$可以視為$E[e^{\eta f(Z)}]$線性內插所得的結果，因此，$E[e^{\eta f(Z)}] \le e^{\eta E[f(Z)]}$
2. $a \le E[f(Z)] \le b$，因此 $\frac{b - E[f(Z)]}{b-a} e^{\eta a} + \frac{E[f(Z)] - a}{b-a} e^{\eta b}$ 可視為$e^{\eta E[f(Z)]}$在exponential函數上線性內插所得的結果，因此， $e^{\eta E[f(Z)]} < \frac{b - E[f(Z)]}{b-a} e^{\eta a} + \frac{E[f(Z)] - a}{b-a} e^{\eta b}$
3. 集1.和2.，故得證

3. Let $\mathcal{H}$ be a function class of VC dimension v. Based on Sauer's lemma, show that for $N > v$,

$$\mathcal{S}_{\mathcal{H}}(N) \leq \left(\frac{eN}{v}\right)^v .$$

- By Sauer's lemma, $S_H(n) \leq \sum_{k=0}^{v} \binom{n}{k}$ for all H and n.

- If we can proof that $\sum_{k=0}^{v} \binom{N}{k} \leq \left(\frac{eN}{v}\right)^v$, we will have $S_H(N) \leq \left(\frac{eN}{v}\right)^v$.

- We can proof $\left(\frac{v}{N}\right)^v \sum_{k=0}^{v} \binom{N}{k} \leq e^v$ instead.

- $\left(\frac{v}{N}\right)^{v} \sum_{k=0}^{v} \binom{N}{k} \le \left(\frac{v}{N}\right)^{k} \sum_{k=0}^{v} \binom{N}{k}$, since $k \le v$,and $v \le N$.
- $\left(\frac{v}{N}\right)^{k} \sum_{k=0}^{v} \binom{N}{k} \le \left(\frac{v}{N}\right)^{k} \sum_{k=0}^{N} \binom{N}{k}$, since $v \le N$.
- By binomial theorem, $\left(\frac{v}{N}\right)^{k} \sum_{k=0}^{N} \binom{N}{k} = \sum_{k=0}^{N} \binom{N}{k} \left(\frac{v}{N}\right)^{k} 1^{N-k} = \left(1+\frac{v}{N}\right)^{N}$.
- $\lim_{N\to\infty} \left(1+\frac{v}{N}\right)^{N} = e^{v}$, and $\left(1+\frac{v}{N}\right) > 1$, so $\left(1+\frac{v}{N}\right)^{N} < e^{v}$.
- Hence, with the properties of the problem, $S_H(n) \le \left(\frac{eN}{v}\right)^{v}$.

4. Show that $Var[1(g^*(\boldsymbol{x}) \neq r)] \leq \frac{1}{4}$.

- $l(g^*(\boldsymbol{X}) \neq r) \in \{ 0, 1 \}$
- 假設有n個樣本，使得 $\Sigma_{i=1}^{n} l(g^*(\boldsymbol{x}_i) \neq r) = np$，其中$0 \leq p \leq 1$，且$np \in Z$
- 那麼，$E[l(g^*(\boldsymbol{X}) \neq r)] = p$，即其平均值。
- $Var[l(g^*(\boldsymbol{X}) \neq r)]$，即變異數為$(\, np \times (1-p)^2 + (n-np) \times (0-p)^2 \,)/n = p-p^2$
- $p-p^2$有極大值在$p=0.5$，此時$Var[l(g^*(\boldsymbol{X}) \neq r)]$為0.25
- 因此，$Var[l(g^*(\boldsymbol{X}) \neq r)] \leq 1/4$