在 1D 中，$d_{vc}=2$；在 2D 中，$d_{vc}=3$，所以我们假设在多维平面中，$d_{vc}=d+1$（d 为维数）

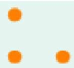
证明:①：$d_{vc}\geqslant d+1$

证明这个不等式的话，我们只要找到一类 d+1 个输入能够被 shatter，就会得到 d_{vc}\geq d+1。因此，只要构造一个 d 维的矩阵 X 能够被 shatter 就行。X 为 d 维矩阵，其中有 d+1 个输入 inputs，每个输入 input 前加上第 0 维度的常数项 1，得到矩阵 X（注意：矩阵 X 存在逆矩阵且唯一）如下：

- some 'trivial' inputs:

$$X = \begin{bmatrix} — \mathbf{x}_1^T — \\ — \mathbf{x}_2^T — \\ — \mathbf{x}_3^T — \\ \vdots \\ — \mathbf{x}_{d+1}^T — \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & 0 \\ 1 & 0 & \ldots & 0 & 1 \end{bmatrix}$$

如果在二维平面上，则相应的矩阵资料代表的点有（0,0）（1,0）（0,1），可以很轻易看出，这三个点是可以 shatter 的。

- visually in 2D: 

下面证明在 d 维度下 d+1 个点是可以被 shatter 的。矩阵 X 可逆的意义是什么？查看 X 是否能被 shatter，对于给定的任意一种 y{+1，-1，..........}，因为矩阵 X 可逆，可以找到一个矩阵 w，$w=X^{-1}y$ 使得(Xw)=y，因此可得 sign(Xw)=y。

to shatter . . .

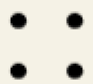
for any $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_{d+1} \end{bmatrix}$, find $\mathbf{w}$ such that

$$\text{sign}(X\mathbf{w}) = \mathbf{y} \quad \Longleftarrow \quad (X\mathbf{w}) = \mathbf{y} \quad \overset{X \text{ invertible!}}{\Longleftrightarrow} \quad \mathbf{w} = X^{-1}\mathbf{y}$$

这个 X 能够被 shatter，就可以得出 $d_{vc}\geqslant d+1$。

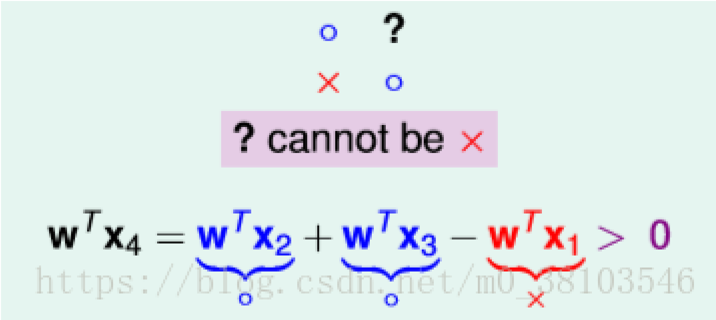
②：$d_{vc}\leqslant d+1$

需要证明任何 d+2 个 input 都不能被 shatter。在二维平面上有 4 个输入的数据

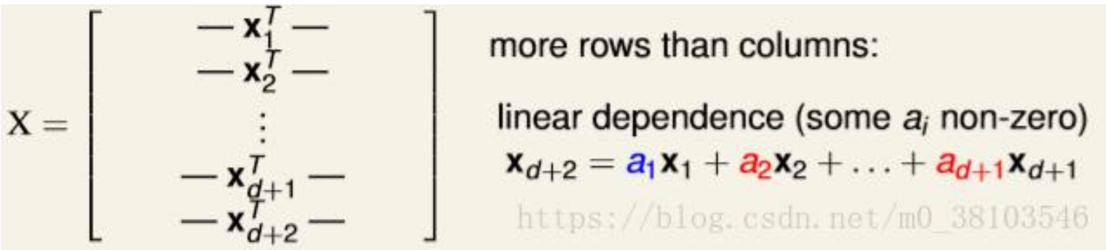


$$X = \begin{bmatrix} — \mathbf{x}_1^T — \\ — \mathbf{x}_2^T — \\ — \mathbf{x}_3^T — \\ — \mathbf{x}_4^T — \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

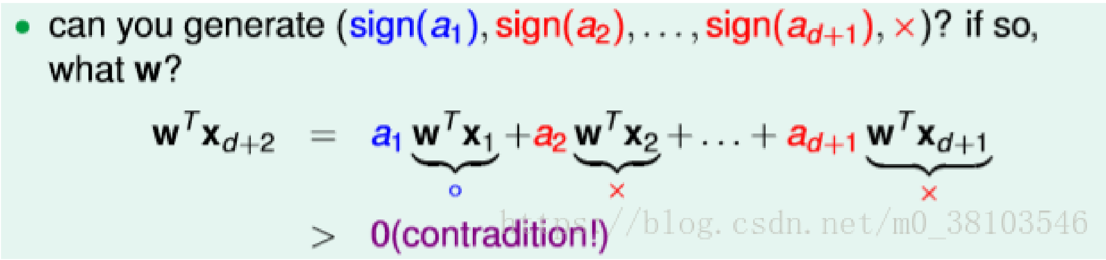
其中有一种情况一定不能被 shatter。



那么对于任意维度 d，查看 d+2 个输入的资料不能 shatter，因此，我们使用一个矩阵 X,其中列的数量为 d+2,行的数量为 d+1（列数大于行数）,这个矩阵不是可逆的，存在一个向量 $X_{d+2}=a_1x_1+a_2x_2+.......+a_{d+1}x_{d+1}$（线性相关，$a_i$ 不会全部都是 0），就不能产生所有的 dichotomy，所以 d+2 个点不能够被 shatter。



如果有一组 w，在 d+1 个输入中产生的情形刚好和 a_{i}的符号一致，因为 x_{d+2} 可以使用其他向量表示，所以得出的结果 w^Tx_{d+2} 一定是正的，即前面 d+1 个决定之后，第 d+2 就必须为正或者负（即确定的），因此无法产生所有的 dichotomy，因此任何 d+2 个输入不能 shatter



问题 1：对于线性的激活函数，

$w_2^T (W_1v + b_1) + b_2 = w_2^T W_1v + (w_2^T b_1 + b_2)$

意味着线性激活神经网络的能力不能超过线性分类器即 VC≤d+1

同时对于第二层 VC≤n+1

所以：VC=min（n+1, d+1）

问题 2：从 d=1 开始，设 n+1 个点是(1), ...,(n),(0)，W1=((1,1,…,1)), b1=(-1, -2,…,-n+1, 0).那么 $W_1v+ b_1$ 为(1, 0, ..., 0),(1, 1, ..., 0), ...,(1, 1, ..., 1),(0, 0, ..., 0)，设 $W_2 = (l_1 - l_{n+1}, l_2 - l_1 - l_{n+1}, l_3 - (l_1 + l_2) - l_{n+1}, ..., l_n - \sum_{i<n} l_i$ l_i)-l_{n+1} 和 b1= l_{n+1},所以 VC≥n+1

同时根据开头证明可得 VC≤n+1，所以 VC=n+1

参考：1.助教给出的参考答案。

2.博客 https://blog.csdn.net/m0_38103546/article/details/82393476

3.周志华《机器学习》第 12 章.计算学习理论
4.台大机器学习基石课程
5. http://www.csuldw.com/2016/08/23/2016-08-23-vc-dimentions/