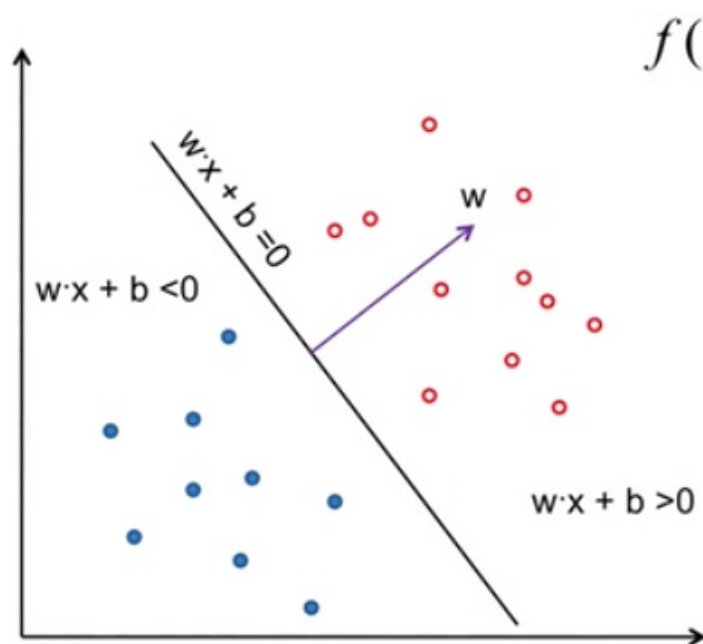
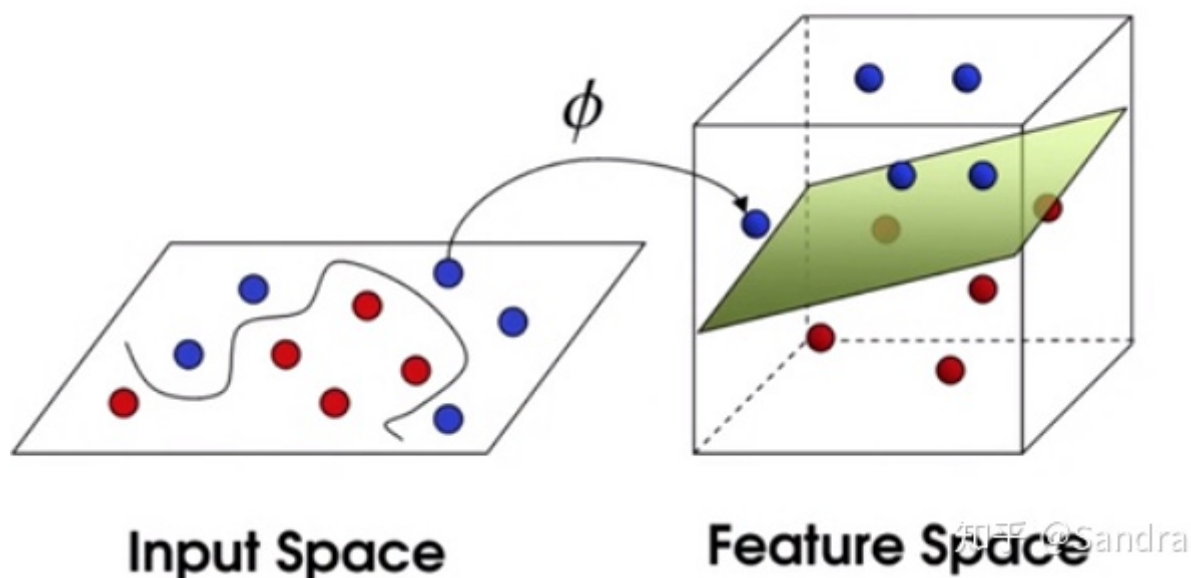


SVM的最核心的思想就是从输入空间 (input space) 向一个更加高维度 (feature space) 的映射。与神经网络的隐含层相似，从输入向某一个中间的阶段做了一个映射，再进行分类。最本源，是一个线性分类器。



$$f(x, w, b) = \text{sign}(g(x))$$

$$= \text{sign}(w \cdot x + b)$$

Just in case ...

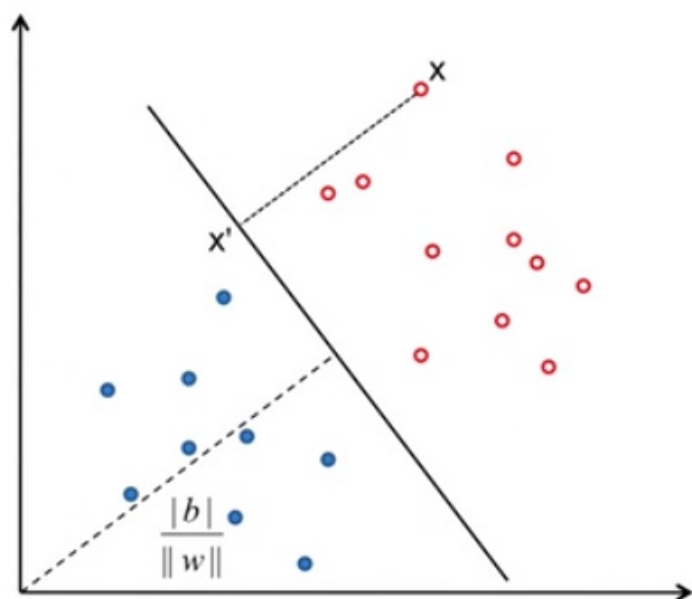
$$w \cdot x = \sum_{i=1}^n w_i x_i$$

$$w \cdot x_1 + b = w \cdot x_2 + b$$

$$w(x_1 - x_2) = 0$$

*w*的方向是与平面垂直的

空间的点到超平面的距离表示方式：



$$g(x) = w \cdot x + b$$

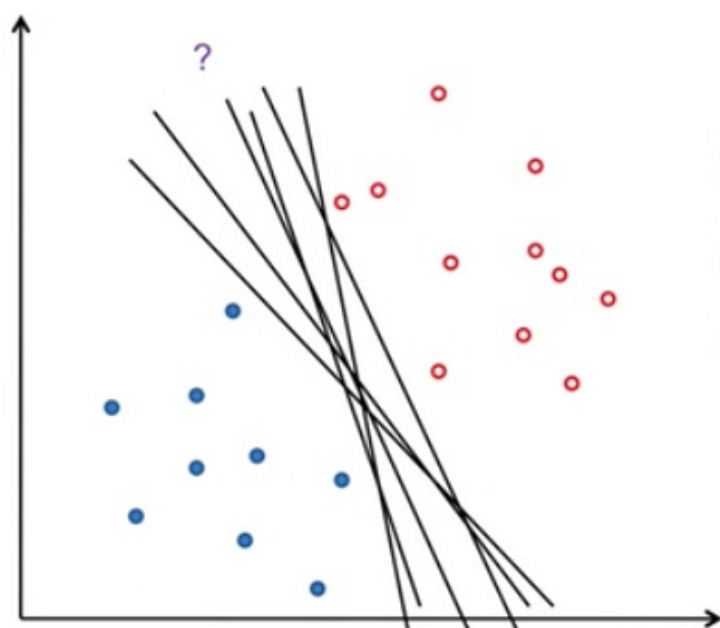
$$x = x' + \lambda w$$

$$\begin{aligned} g(x) &= w(x' + \lambda \cdot w) + b \\ &= w \cdot x' + b + \lambda w \cdot w \\ &= \lambda w \cdot w \end{aligned}$$

$$\begin{aligned} M &= \|x - x'\| = \|\lambda w\| \\ &= \frac{|g(x)| \times \|w\|}{w \cdot w} = \frac{|g(x)|}{\|w\|} \end{aligned}$$

点到平面距离公式M要记住，后面会应用

选择最优的平面：



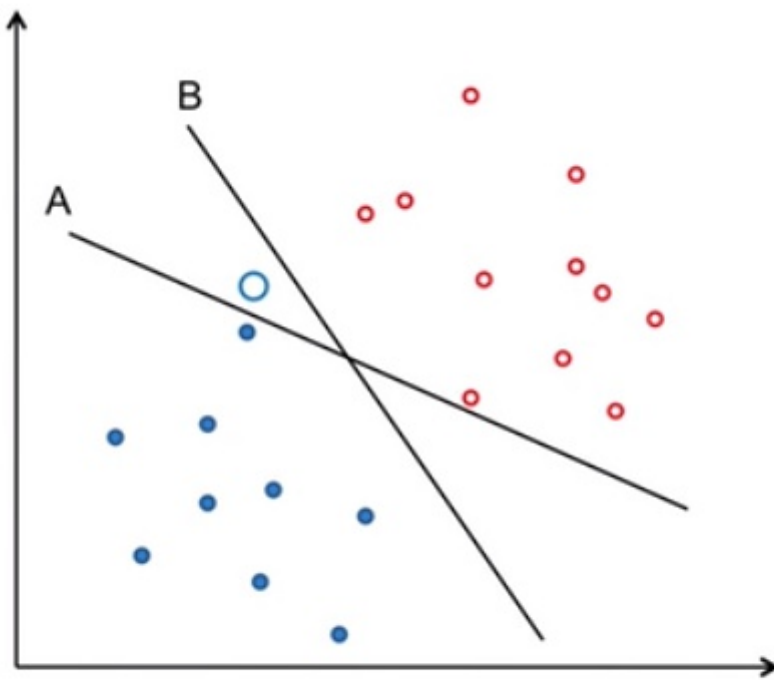
Which classifier is the best?

All have the same training error.

How about generalization?

知乎 @Sandra

中间的空心大圆圈是一个未知样本，根据不同的分界面，它被分为了不同的类型。那哪一种更好呢？

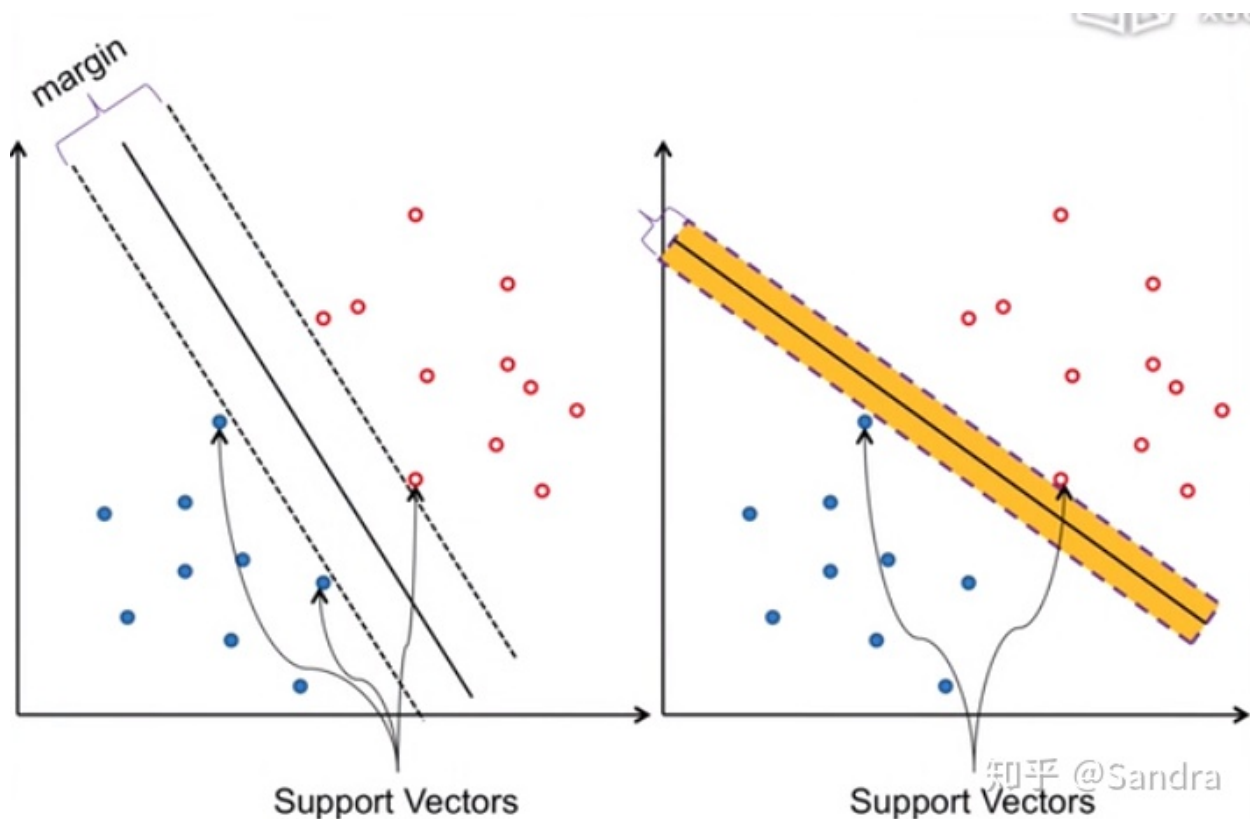


Classifier B divides the space more consistently (unbiased).

B比A好，因为B将空间分割的更加无偏。因为中间的空白区域，所以A,B都看似可行并且不相同，所以无法确定分界面的方向。B基本将其均分开，A靠着蓝点和红点都很近，有很大的bias。

Margin:

将分界面向两个方向平移至不能平移的位置（他碰到了点），可以平移的距离叫做Margin（间隔）。正好卡住这些分界面的点称为Support Vectors。不同方向的Margin不同，Support Vectors也不同。



Support Vectors (很少的点) 决定了分界面移动的范围 (Margin)

直观上说, Margin越大, 容错性越强。所以, 希望这个分界面的Margin越大越好。SVM就可以最大化Margin (线性支持向量机)。

- ❖ The margin of a linear classifier is defined as the width that the boundary could be increased by before hitting a data point.
- ❖ Intuitively, it is safer to choose a classifier with a larger margin.
- ❖ Wider buffer zone for mistakes
- ❖ The hyperplane is decided by only a few data points.
 - Support Vectors
 - Others can be discarded!
- ❖ Select the classifier with the maximum margin.
 - Linear Support Vector Machines (LSVM)
- ❖ How to specify the margin formally?

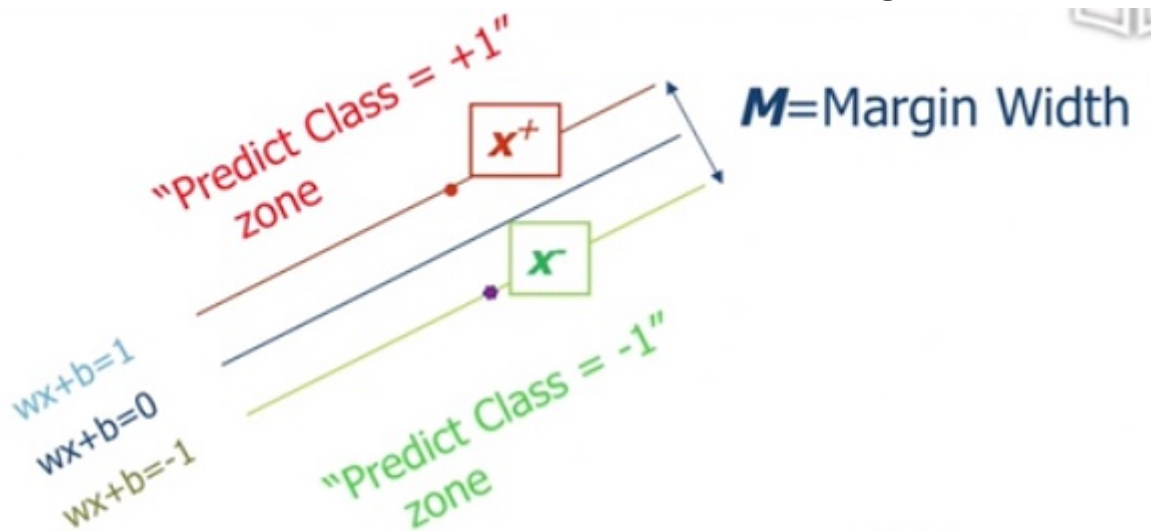


只有很少的对分界面有影响

Margin如何定义:

上面的线定义为 $=1$ ，下面的线定义为 $=-1$ （也可以令其等于正负一，因为可以进行缩放）

--> 也就是说在这个分界面上方都定义为1，下方都定义为-1（用正负一来定义点的两个类型）。运用上面点到平面的公式，得到Margin的式子。



$$M = \frac{2}{\|w\|}$$



知乎 @Sandra

中间部分是空的，没有点的，所以点要么在上面，要么在下面。