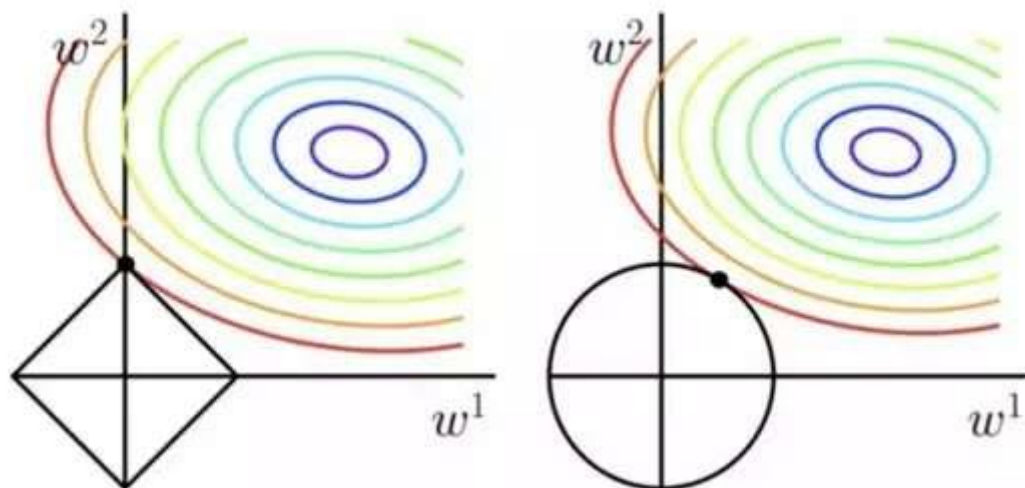


正则化：

正则化是为了解决过拟合问题。目标函数中加入正则化，即加入模型复杂度的评估。正则化符合奥卡姆剃刀原理，即：在所有可能的模型中能够很好的解释已知数据并且十分简单的模型才是最好的模型。

L1 正则化：向量中各元素绝对值的和。关键在于能够对特征进行自动选择，系数参数可以减少非必要的特征引入噪声。

L2 正则化：向量中各元素的平方和。L2 会使得各元素尽可能小，但都不为零。



左边为 L1 正则化，右边为 L2 正则化。假设权重参数  $w$  只有二维  $w_1$  和  $w_2$ 。L1 为各元素绝对值和，即得到的形状为菱形，L2 形状为圆。很容易可以发现 L1 更容易在顶点处相切，L2 则不容易在顶点处相切。顶点处则其中一个参数为零，这就是为什么 L1 会使得参数稀疏的原因。