

图 3-7 二维平面上线性不可分样本的例子

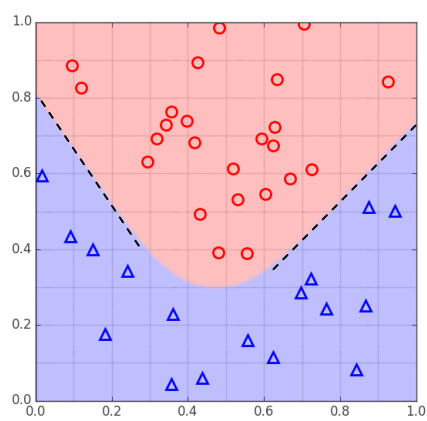


图 3-9 样本和二分类模型的分类边界

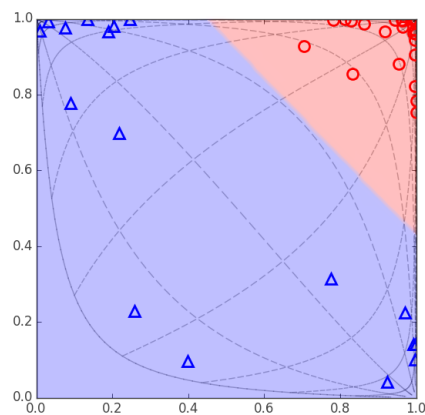


图 3-10 输入样本经过隐藏层非线性变换后在新的二维空间中的坐标

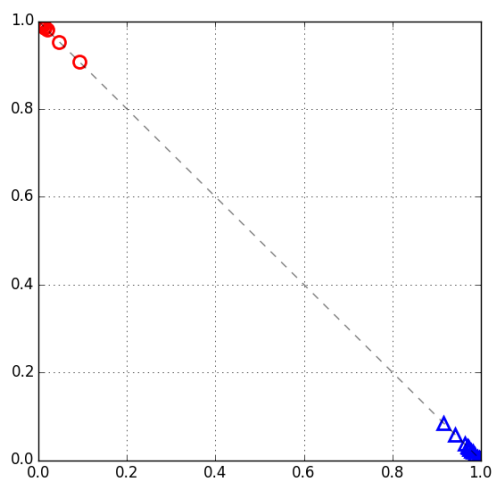


图 3-11 Softmax 层输出的结果表示在二维平面中

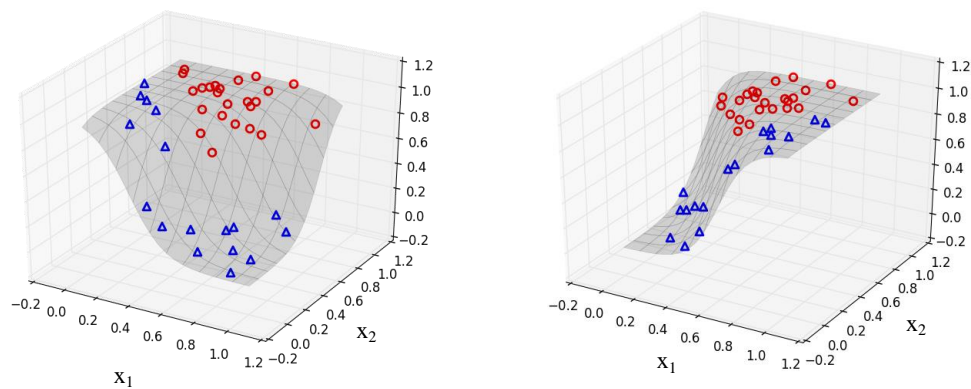
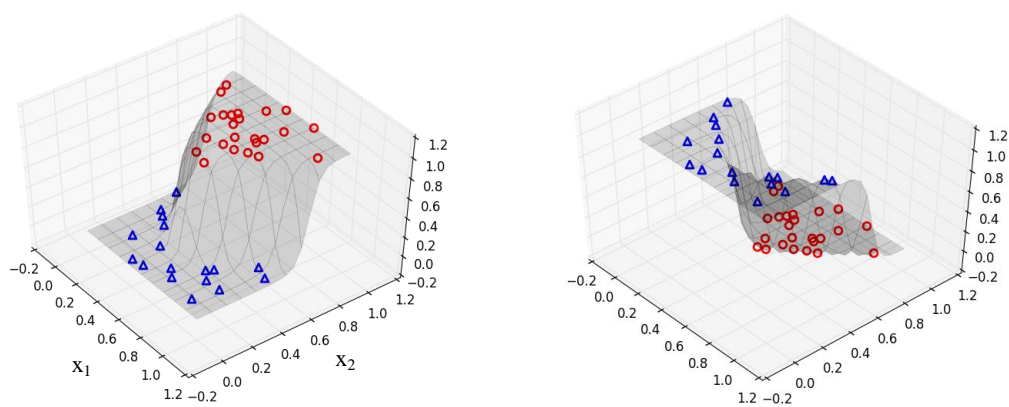
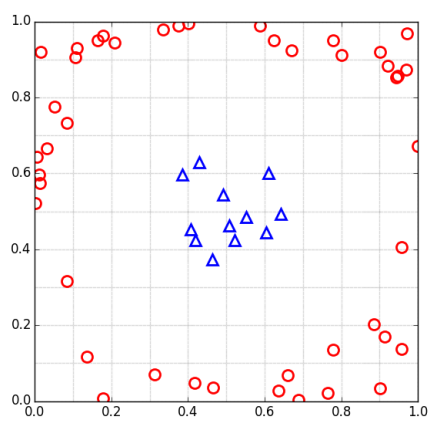
图 3-12 将隐藏层的输出作为 z 轴画出的样本在 3 维空间中的分布图 3-14 将 Softmax 输出作为 z 轴，样本在三维空间中的分布

图 3-15 一个两隐层单元网络不可分的样本在二维平面的分布

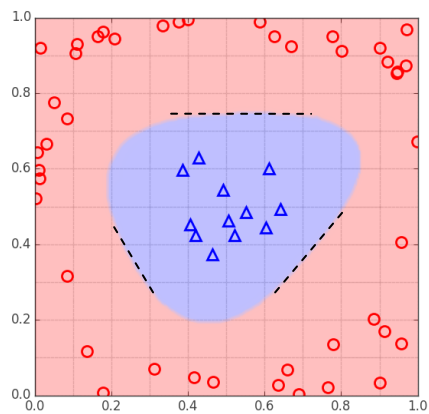


图 3-16 用三个隐层单元后得到的分类结果

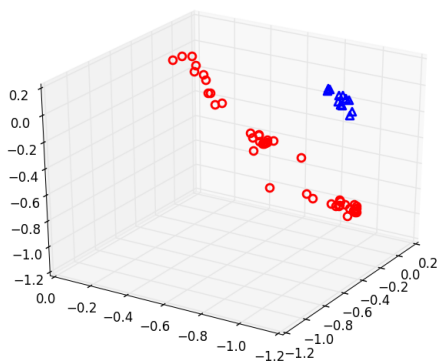


图 3-17 图 3-16 的样本在隐层对应的三维空间中的分步

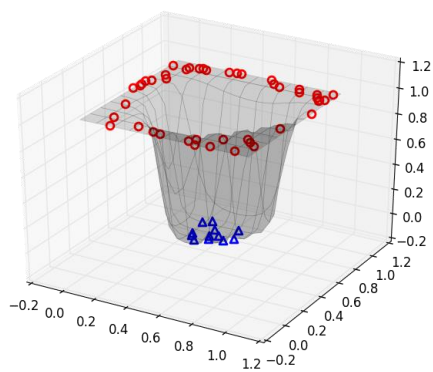


图 3-18 将 Softmax 输出作为 z 轴，样本在三维空间中的分布

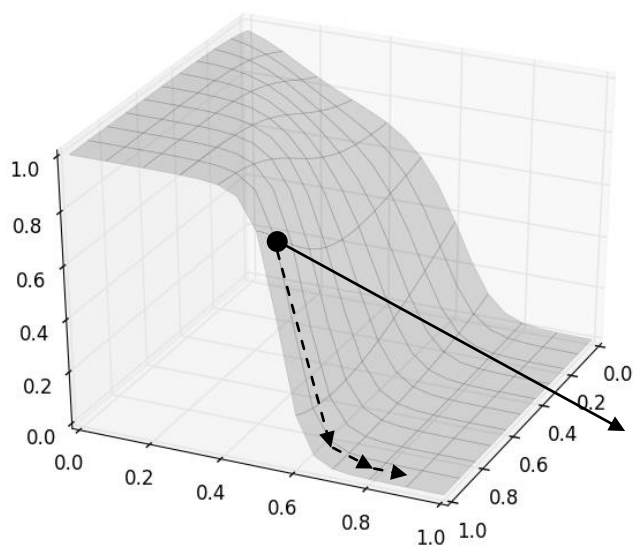


图 3-28 “断崖”和梯度爆炸

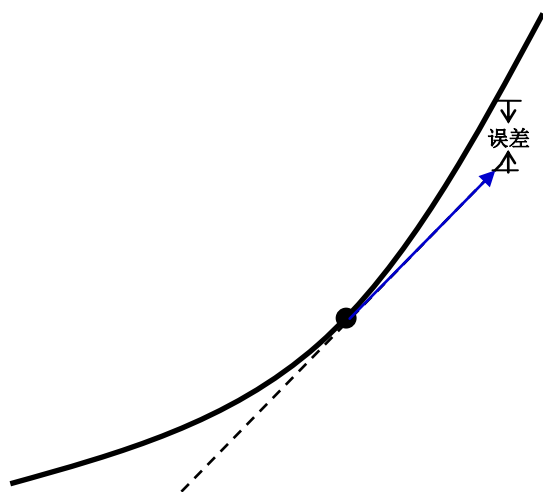
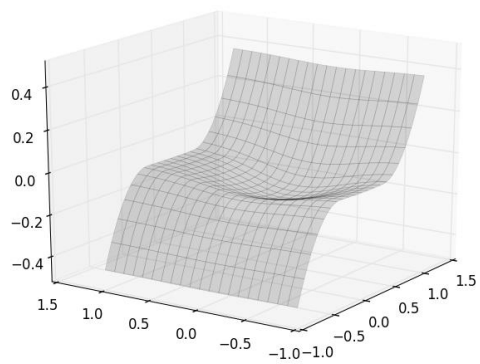
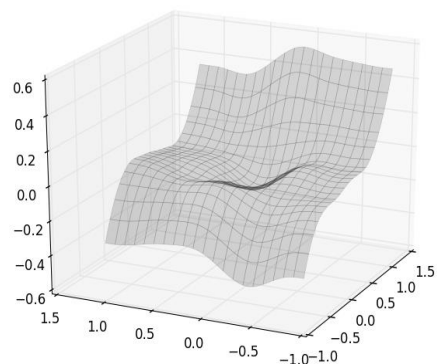


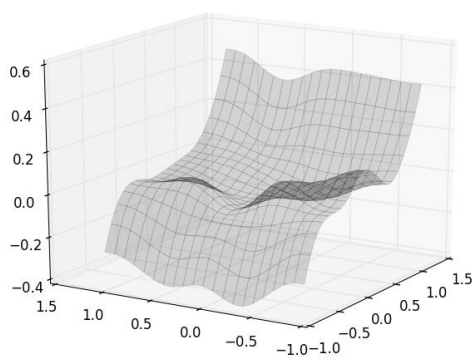
图 3-29 梯度和线性近似



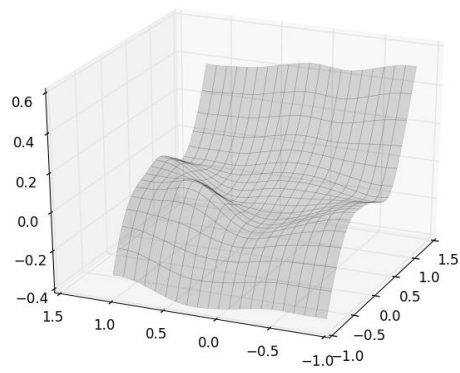
a



b



c



d

图 3-30 全量梯度下降和 SGD 对应的损失函数曲面

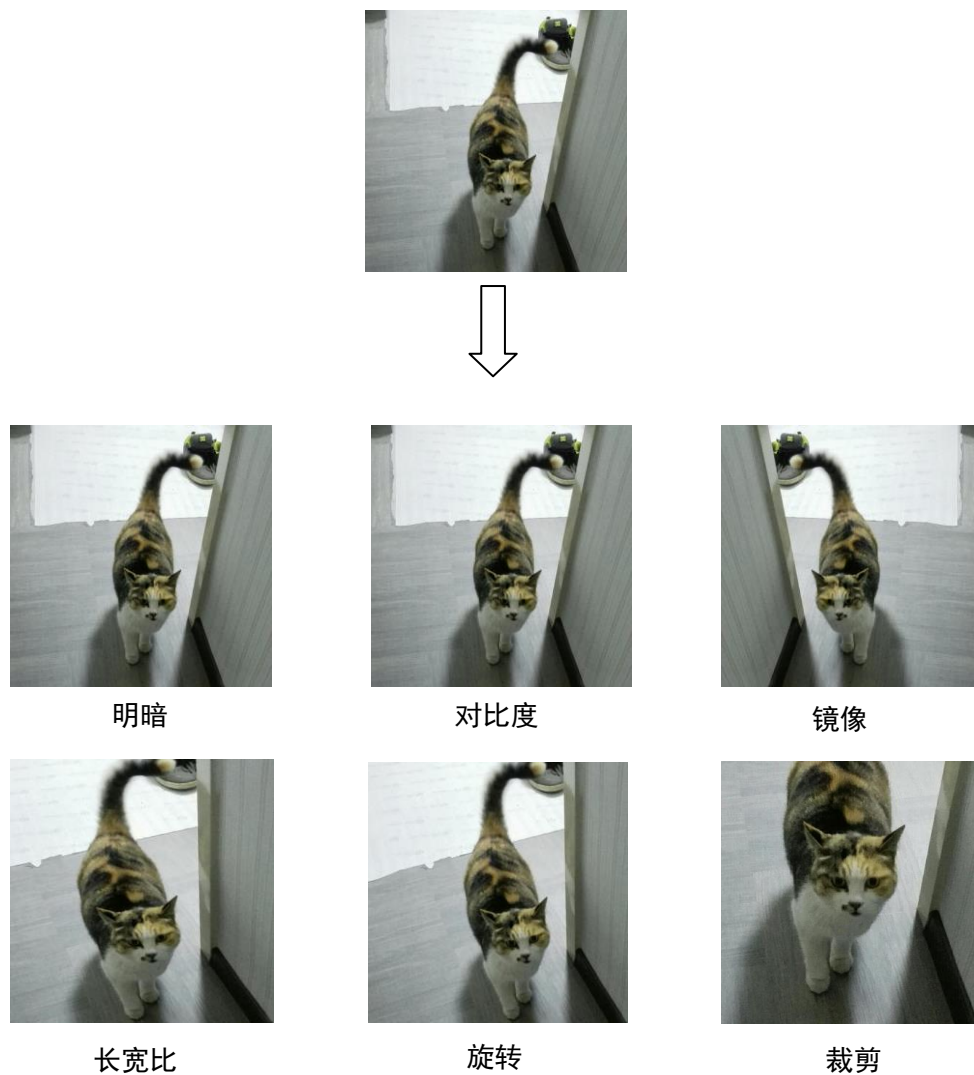


图 3-31 猫图片数据增加的例子

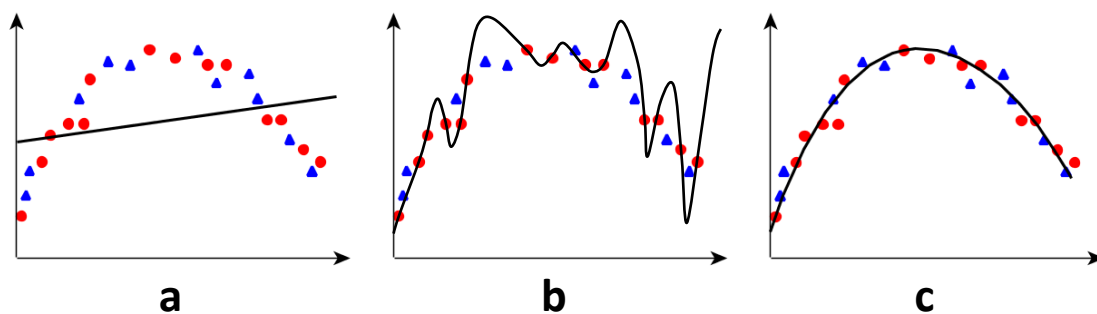


图 3-32 一维拟合的例子

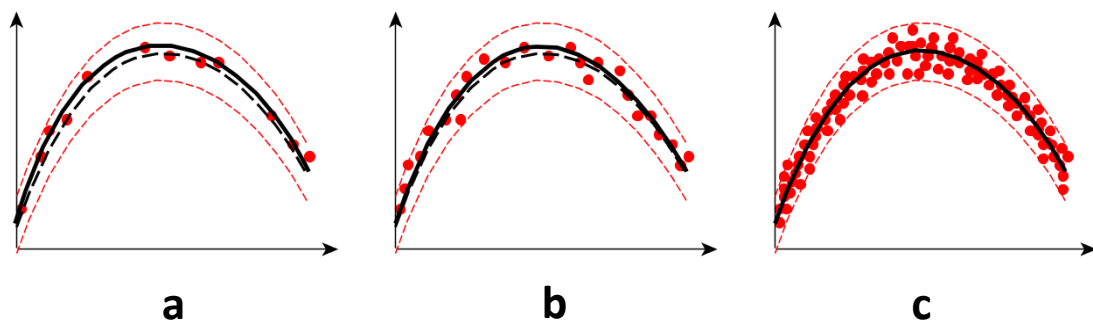


图 3-34 完美模型的误差随数据量的变化

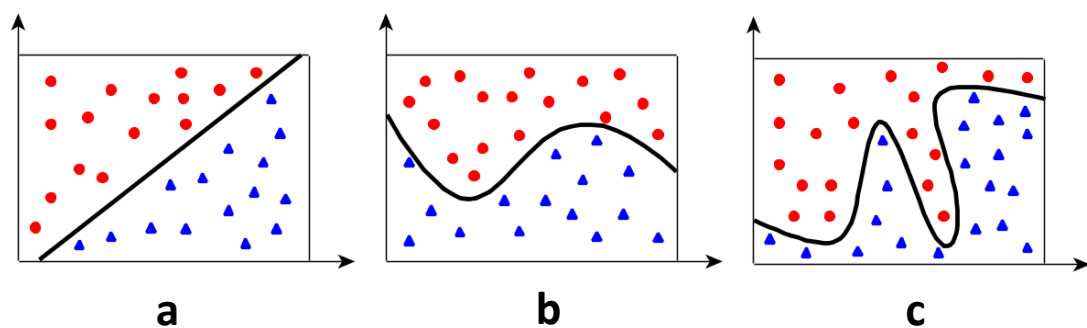


图 3-35 没有对任何分布都有效的分类算法(NFL)

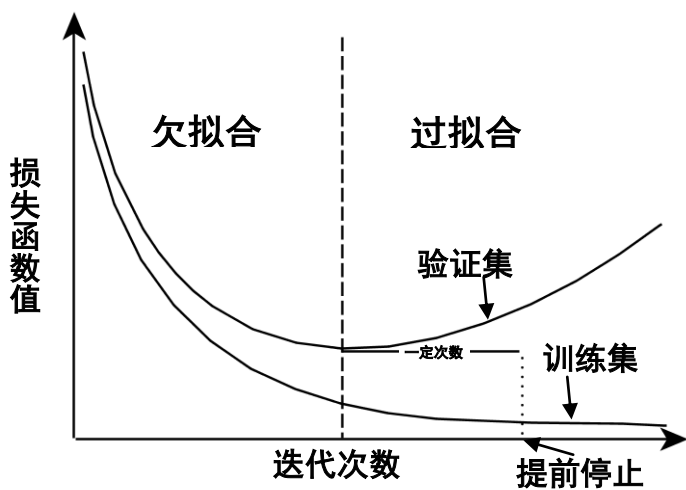


图 3-36 训练过程中损失函数在验证集和训练集上的趋势

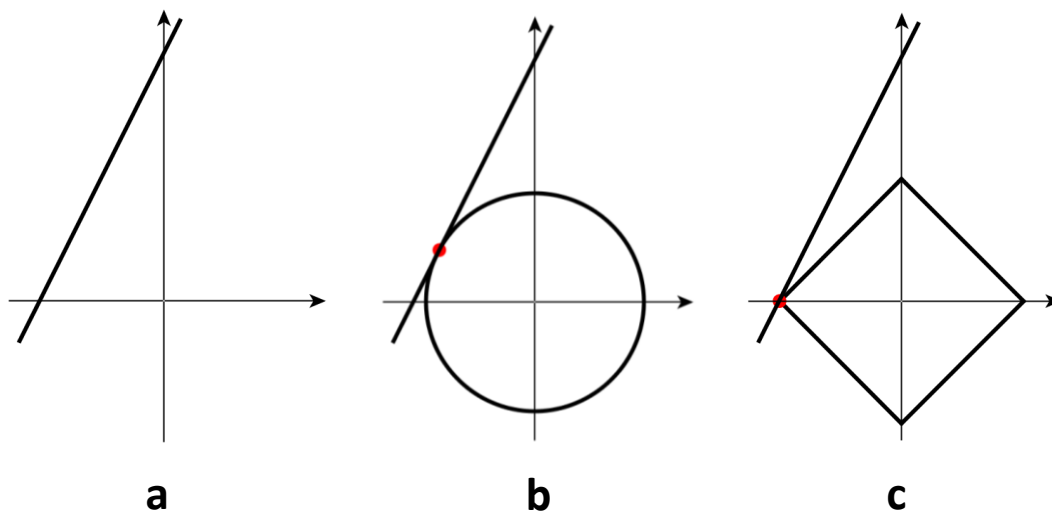


图 3-37 病态方程求解和规范化

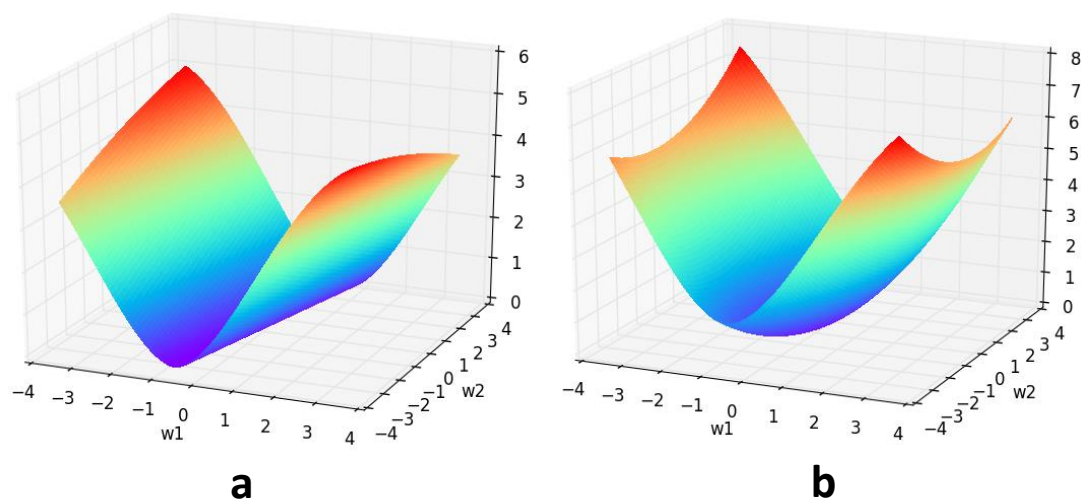


图 3-38 L2 规范化对目标函数曲面的影响

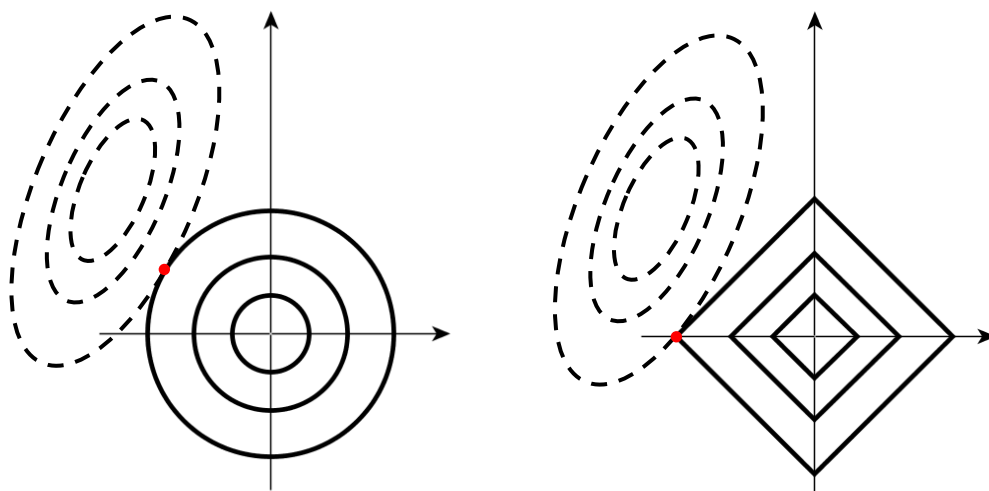


图 3-39 L2 规范化和 L1 规范化的区别

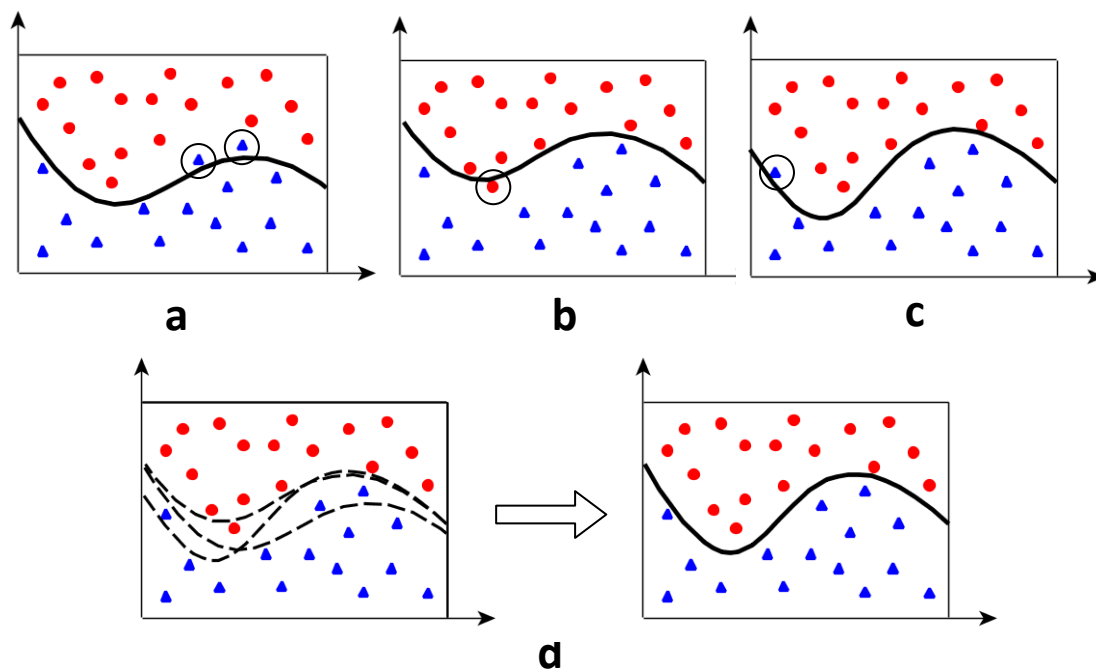


图 3-40 集成 (ensemble) 的示意例子

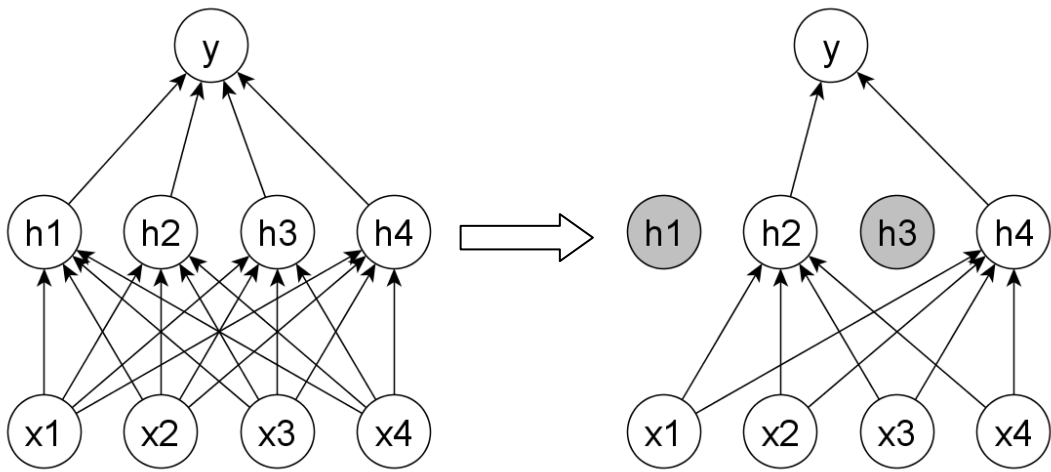


图 3-41 随机失活 (dropout) 示意

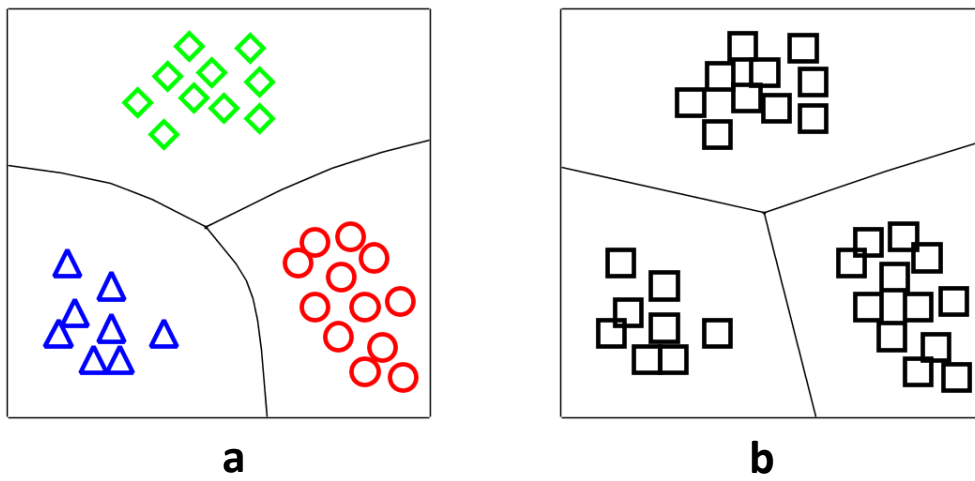


图 3-42 监督学习和无监督学习示意