1. 线性单元（解决回归问题：返回实数值）

当面对的数据集不是线性可分的（分开后有交集），感知器规则可能无法收敛->无法完成训练

解决：使用可导的线性函数来替代感知器的阶跃函数->线性单元（能收敛）

例子：（思想）在处理语音识别的时候，获取语音容易，但是一个个标注相关文本较为复杂，所以先用**无监督方法做一些聚类，模型总结相似音节，用少量带标注的训练样本告诉模型这些音节对应的文字**

1. 目标函数

监督学习下：x表示特征，y表示实际值(label)，表示预测值

其中y与 的接近程度，数学上有很多的方法：

如用e = 1/2(y - 表示单个样本的误差

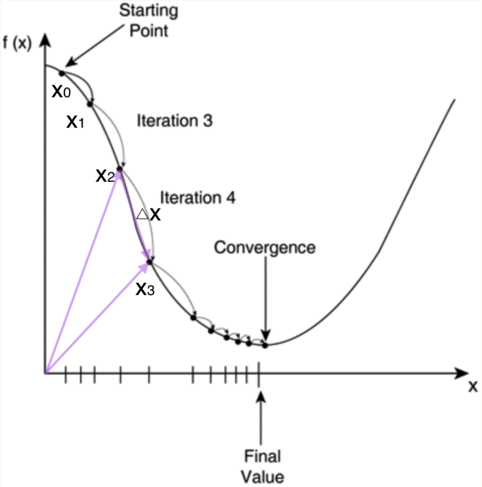
用误差和表示模型的误差：

= =

由此可见，模型的训练=求取w使得式子取最小值

1. 梯度下降优化算法

对于计算机：可以凭借计算能力，一步步将函数的极值点试出来



任意选定一点，如 ，每次迭代修改x的值

**梯度：指向函数数值上升最快的方向（函数定义：相对于各个变量的偏导数）**

梯度下降算法的公式：

是的梯度，是步长即学习速率

则结合上文的目标函数，梯度下降算法可写为

梯度上升算法为：

进一步得到：