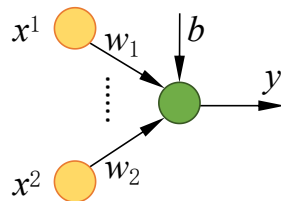
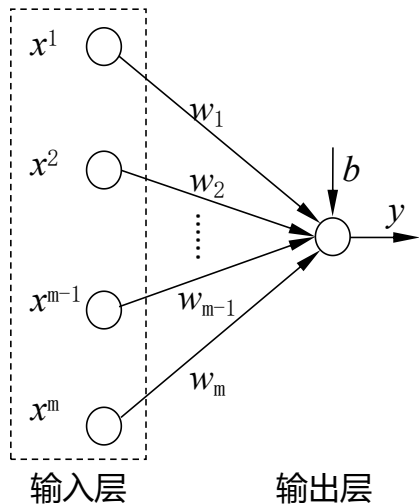




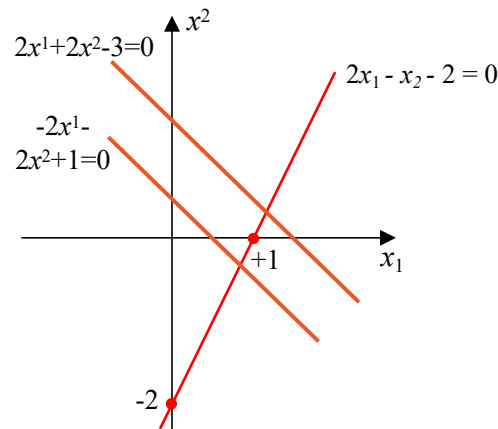
12.3.2 超参数和验证集

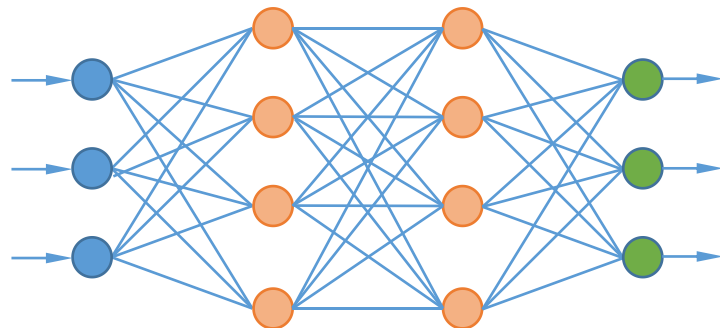
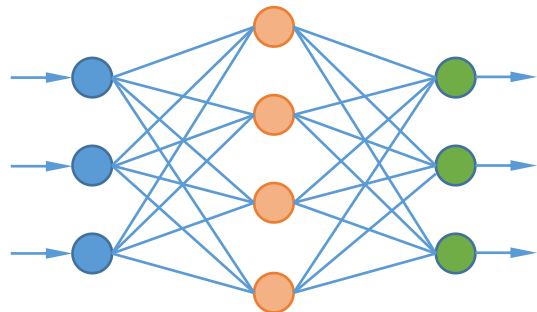
感知机



$$y = \sigma(z) = \sigma(w_1 x^1 + w_2 x^2 + b)$$

w_1, w_2, b	线性分类器
2, -1, 2	$y = \sigma(2x^1 - x_2 - 2)$
2, 2, -3	$y = \sigma(2x^1 + 2x^2 - 3)$ AND
-2, -2, 1	$y = \sigma(-2x^1 - 2x^2 + 1)$ NOR





□ 万能近似定

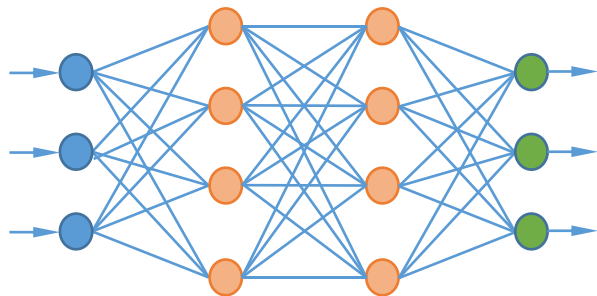
在**前馈型神经网络**中，只要有一个**隐含层**，并且这个隐含层中有**足够多的神经元**，就可以逼近任意一个**连续的**函数或空间分布

□ 多隐含层神经网络

- 能够表示**非连续的**函数或空间区域
- 减少泛化误差
- 减少每层神经元的数量



□ 隐含层的设计



通过实验来确定超参数

layer $\in (1, 2, 3)$

node $\in (4, 16, 32, 64)$

可能的组合: $3 \times 4 = 12$ 种

- 使用**训练集**训练好所有模型
- 使用**同一个**测试集, 得到不同模型上的误差
- 选出测试集上**误差最小**的模型



□ 训练集、验证集和测试集

训练集：训练模型，确定模型参数

验证集：确定超参数

测试集：评估模型的泛化能力

