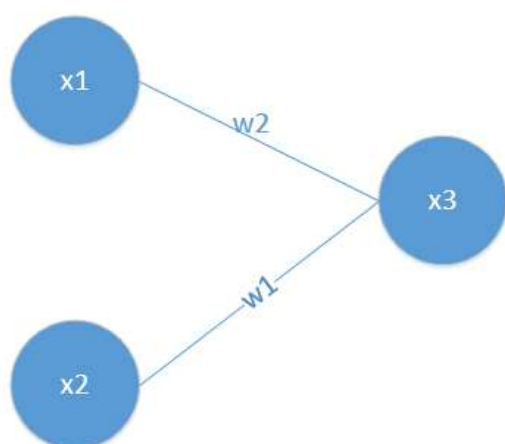


# 特征组合的角度理解激活函数

激活函数其中一个重要的作用是加入非线性因素的，解决线性模型所不能解决的问题。

下面从特征的组合角度来解释一下激活函数的作用。

一个简单的感知机如下：



其中 $x_1, x_2$ 输入均为特征的输入

$$x_3 = w_1 * x_1 + w_2 * x_2$$

激活函数采取sigmoid函数,公式表达如下：

$$S(x) = \frac{1}{1 + e^{-x}}$$

$$\text{则 } S(x_3) = \frac{1}{1 + e^{-x_3}}$$

此时，我们可能看不出什么不同，但是根据泰勒展开，

$$e^x = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + o(x^3)$$

我们能够看到，将 $x_3$ 代码到激活函数的时候，其实激活函数的 $e^x$ 泰勒展开后，有平方项，有立方项，有更高的项，**这些自动能够把输入特征进行两两组合，进行三三组合或者其它的组合。**

**比如其中的平方项体现了将特征进行两两组合：**

$$(w1 * x1 + w2 * x2)^2 = (w1 * w1 * x1 * x1 + w2 * w2 * x2 * x2 + w1 * x1 * x2 * w2)$$

**这就把原来需要领域知识的专家对特征进行组合的情况，在激活函数运算后，其实也能够起到类似特征组合的作用。**

**(只要激活函数中有能够泰勒展开的函数，就可能起到特征组合的作用)**