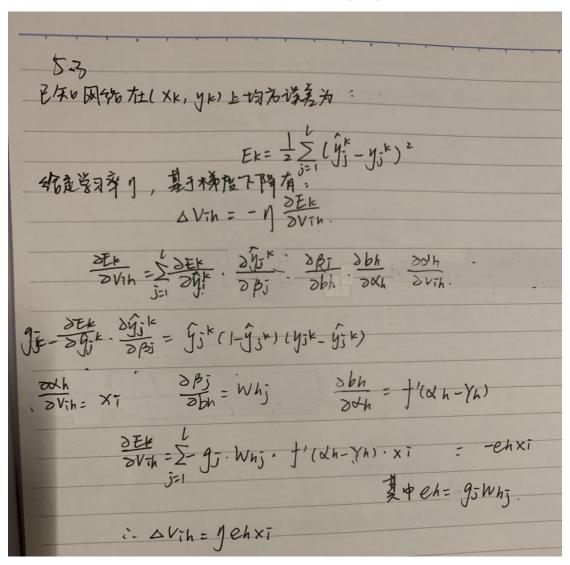
## 5.1 试述将线性函数 $f(x) = w^T x$ 用作神经元激活函数的缺陷?

如果激活函数是一个线性函数,那么无论多少层网络,都表示为一层线性网络。如果用题目 f(x) 做激活函数,无论多少层神经网络都退化成了线性回归,达不到"激活"与"筛选"的目的。

## 5.2 试述使用图 5.2(b)激活函数的神经元与对率回归的联系

对率回归,是使用 Sigmoid 函数作为联系函数时的广义线性模型。单位阶跃函数不连续,难以求导,所以用对数几率函数代替。使用 Sigmoid 激活函数,每个神经元几乎和对率回归相同,只不过对率回归在 [sigmoid(x)>0.5] 时输出为 1, 而神经元直接输出 [sigmoid(x)] 。

5.3 对于图 5.7 中的 Vh,试推导出 BP 算法中的更新公式(5.13)



## 5.4 试述式(5.6)中的学习率的取值对神经网络训练的影响

学习率  $\eta \in (0,1)$ 控制算法每一轮迭代中的更新步长,学习率  $\eta$ 太小,则每次的更新量会太小,这使得迭代次数增多;反之,如果学习率  $\eta$ 太大,则会出现震荡情况,即在最小值附近来回波动,导致算法无法收敛。