累积BP

将单个 E_k 误差相加，则对 w_{hj} 求偏导也是各项的偏导相加.

$$w_{hj} \leftarrow w_{hj} + \frac{1}{n}\Delta w_{hj}$$

方差与偏差的平衡：使用正则化减小方差，避免过拟合 (overfitting).

$$E = \lambda \frac{1}{m}\sum_{k=1}^{m} E_k + (1-\lambda)\sum_i w_i^2$$

（连接权和阈值，也就是也可用 n^2，参数值 参数）

BP算法会导到局部最小，为了得到比某局部最小更优的解，可以：

① 以不同的参数值初始化多个神经网络

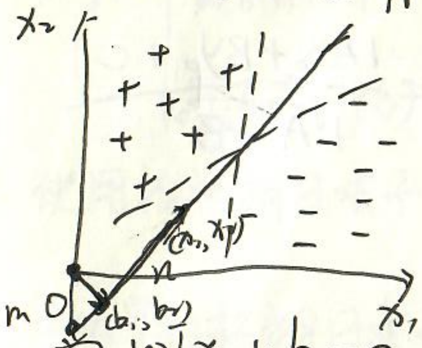
② "模拟退火" 以一定的概率接受比当前解更差的结果（保留每个最小参数空间，限定迭代次数）

③ 随机梯度下降，即使在局部最小点，它计算的梯度仍可能不为零.

增加隐层数目提高拟合能力.

2019.10.04

I 支持向量机 (Support vector machine).



分割两类点的超平面

$$w^T x + b = 0$$

$$w = (w_1; w_2; w_3; \cdots; w_d)$$

当 $w^T x + b > 0$，点在 $w^T x + b$ 的上方

当 $w^T x + b < 0$，点在 $w^T x + b$ 的下方.

w 为超平面 $w^T x + b = 0$ 的法向量

由于向量 x 可分解为 m 与 n，$x = m + n$.

$$w^T(m+n) + b = 0$$

$$w^T m + w^T n + b = 0$$

由于 w 与 n 垂直，$w^T n = 0$

$$w^T m + b = 0, \quad w^T m = -b.$$

样本空间任意点 x 到超平面 (w, b) 距离：

$$r = \frac{|w^T x + b|}{\|w\|}$$

$$[w_1, w_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + b$$

$d = |c|\cos\theta$ =

L_0 范数是指向量中非0的元素个数.

L_2 范数是指向量各元素平方和然后求平方根

L_1 范数是向量各元素之和