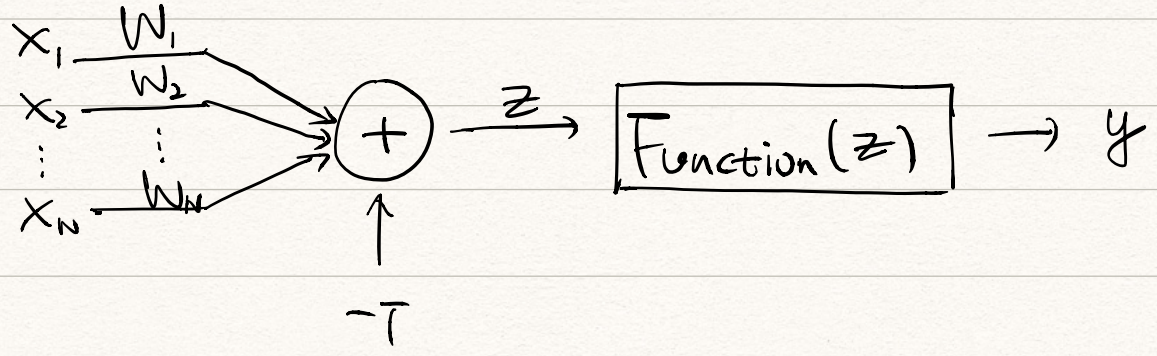
Linear: $f(ax+by) = af(x) + bf(y)$

Affine: $f(x) = \sum_i w_i x_i + b$

The "soft" perceptron (logistic):




$$z = \sum_i w_i x_i - T$$

"affine combination"
仿射组合

translated linear hyperplane

| | | |
|---|---|---|
| sigmoid: | $y = \frac{1}{1+e^{-z}}$ | $y' = y(1-y)$ |
| softplus: | $y = \ln(1+e^z)$ | $y' = \frac{1}{1+e^{-z}}$ |
| Relu: | $y = \max(0, z)$ | $y' = \begin{cases} 1 & z>0 \\ 0 & z\le 0 \end{cases}$ |
| tanh: | $y = \frac{e^z - e^{-z}}{e^z + e^{-z}}$ | $y' = 1 - y^2$ |

Define "depth"

"depth" is the length of the longest path from a source to a sink

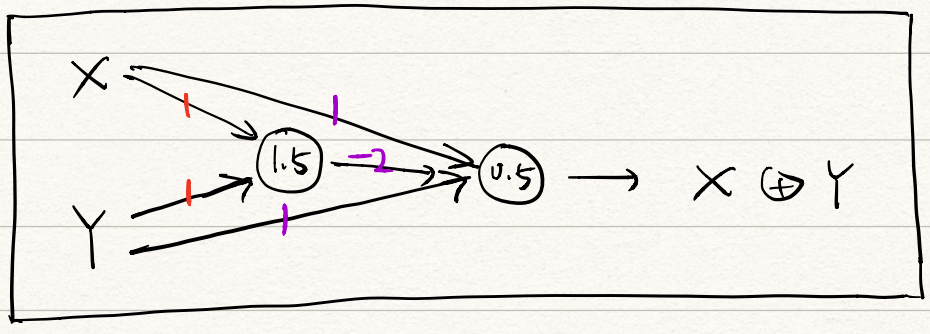
X: A one-hidden-layer MLP is a Universal Boolean Function.

Karnaugh Map (k-map)

$2^{N-1}$ perceptrons in Single hidden layer

$= \frac{2^N}{2}$

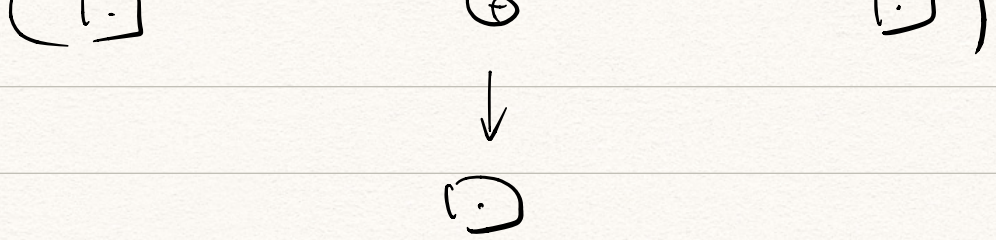
总共 $2^{N-1}+1$ 个 perceptron




异或: 3(N-1) perceptrons for N variables

在 $2\log_2(N)$ layers 中.

$(X_1 \oplus X_2) \oplus (X_3 \oplus X_4) \oplus (X_5 \oplus X_6) \oplus (X_7 \oplus X_8)$



sufficient capacity → represent any function