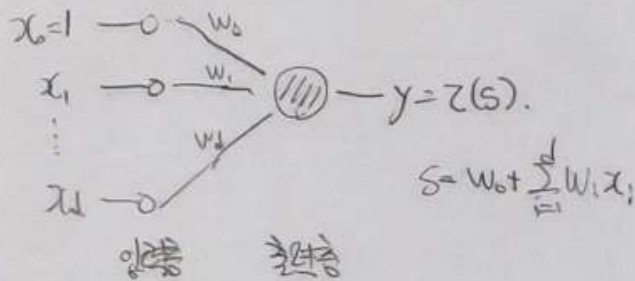


* 기계학습. 6주차 A4 (page 장). 20160409 컴퓨터공학과 김성수

* 퍼셉트론의 구조와 동작.

- 입력층의 i번째 노드 특성벡터 $x = (x_1, x_2, \dots, x_d)^T$ 의 i번째 요소 x_i 를 담당
- 출력층은 한 개의 노드
- i번째 입력 노드와 출력층을 연결하는 Edge는 가중치 w_i 를 가짐
- 해당하는 특징값과 가중치 곱 결과를 더하여 s 구하고, 활성화 함수를 적용함



* Decision Boundary

- 전체 공간을 +1, -1 두 개의 범위로 공간을 분할
- $d(x) = w_1 x_1 + w_2 x_2 + \dots + w_d x_d + w_0 = 0$

* Cost Function

Param 가중치 $w = (w_0, w_1, \dots, w_d)^T$

- ① $J(w) \geq 0$.
- ② w 가 최적이면 $J(w) = 0$
- ③ 틀린 샘플이 많은 w 일수록 $J(w)$ 는 큰 값

Cost Function for Perceptron

$$J(w) = \sum_{x_k \in Y} -y_k (w^T x_k)$$

\rightarrow ① $y_k = 1$ (class) : $\sum_{x_k \in Y} -w^T x_k$
 \rightarrow ② $y_k = -1$ (class) : $\sum_{x_k \in Y} w^T x_k$

* Gradient

$$\frac{\partial J(w)}{\partial w_i} = \sum_{x_k \in Y} -y_k x_{ki}, \quad i=0, 1, \dots, d$$

$$w_i = w_i + \rho \sum_{x_k \in Y} y_k x_{ki}, \quad i=0, 1, \dots, d$$

* Batch vs Stochastic

Batch mode	Stochastic mode.
$Y = \emptyset$	X 의 샘플을 shuffle.
for $j=1$ to n	quit = true
$y = z(w^T x_j)$	for $j=1$ to n
if $(y \neq y_j)$ $Y = Y \cup x_j$	$y = z(w^T x_j)$
if $(Y \neq \emptyset)$	if $(y \neq y_j)$
for $i=0$ to d	quit = false
$w_i = w_i + \rho \sum_{x_k \in Y} y_k x_{ki}$	for $i=0$ to d
until $(Y = \emptyset)$	$w_i = w_i + \rho y_j x_{ji}$
$\hat{w} = w$	until C
	$\hat{w} = w$

* MLP

① Hidden Layer

\rightarrow 원래 특징공간을 높이에 유한함

② Sigmoid function.

\rightarrow 출력값을 신경망으로 간해 유용성 있는 의사결정

③ Back Propagation.

\rightarrow Target과 Output의 차이를 뒤로 전파하여 백스캐싱

* 은닉 노드, 출력 노드의 연산

• 은닉

$$z_j = z(\sum u_j)$$

$$z_{sum_j} = u_j^T x \quad u_j^T = (u_{j0}, u_{j1}, \dots, u_{jd})$$

• 출력

$$x = (1, x_1, x_2, \dots, x_d)^T$$

$$o_k = z(\sum v_k)$$

$$o_{sum_k} = v_k^T z \quad v_k^T = (v_{k0}, \dots, v_{kp})$$

$$z = (1, z_1, \dots, z_p)^T$$