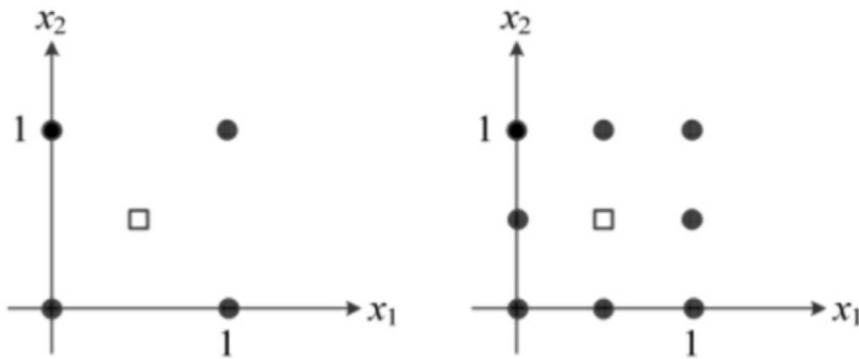


- 8 다음 훈련집합을 3차원 공간에 그리시오. 이 데이터를 최소 오류율로 분류하는 퍼셉트론을 제시하시오.

$$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{x}_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{x}_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \mathbf{x}_6 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$y_1 = 1, \quad y_2 = -1, \quad y_3 = 1, \quad y_4 = -1, \quad y_5 = 1, \quad y_6 = -1$$

- 12 다음 분류 문제에 답하시오.



- (1) 이 분류 문제를 해결하는 다층 퍼셉트론을 각각 제시하시오.
  - (2) 은닉 노드를 2개만 가진 퍼셉트론으로 오른쪽 상황을 해결할 수 있는지 답하시오. 가능하다면 은닉 노드 2개인 다층 퍼셉트론을 채택하는 것이 유리한지를 일반화 능력 측면에서 판단하시오.
- Hint** [그림 11-6]의 여백 개념을 사용하시오.

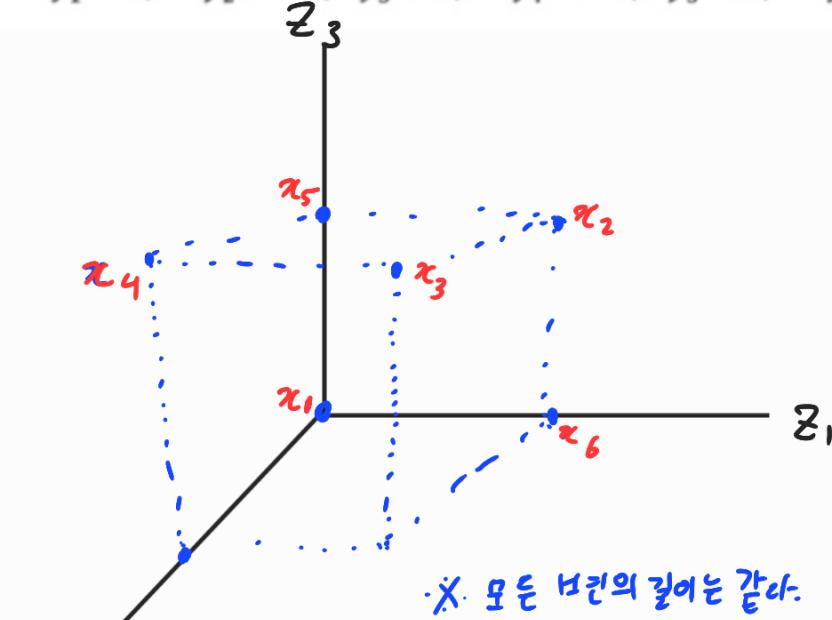
- 13 [알고리즘 3-4]의 부연 설명에서 “ $\mathbb{X}$ 에서 임의로 샘플 하나를 뽑는” 랜덤 샘플링 버전을 소개하였다. 랜덤 샘플링 버전이 MNIST 훈련집합을 사용한다고 가정하자. MNIST는 숫자 부류별로 6천 샘플을 가진다. [알고리즘 3-4]의 라인 2의 루프를 60만 번 반복한다고 가정한다. 10세대를 수행하는 셈이며 각 샘플은 평균 10번 선택된다. 다음 질문에 답하시오.

- (1) 어떤 샘플이 한 번도 선택되지 못할 확률은?
- (2) 어떤 부류에 속한 모든 샘플이 한 번도 선택되지 못할 확률은?
- (3) (1)과 (2)의 분석 결과를 토대로 랜덤 샘플링 방법의 유효성에 대한 의견을 제시하시오.

8 다음 훈련집합을 3차원 공간에 그리시오. 이 데이터를 최소 오류율로 분류하는 퍼셉트론을 제시하시오.

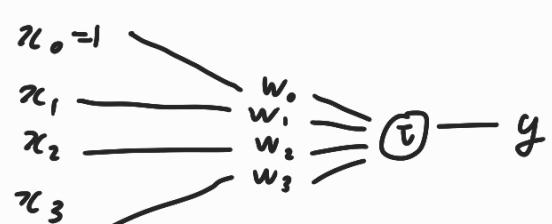
$$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{x}_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{x}_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \mathbf{x}_6 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$y_1 = 1, \quad y_2 = -1, \quad y_3 = 1, \quad y_4 = -1, \quad y_5 = 1, \quad y_6 = -1$$



$$\tau(s) = \text{step function} = \begin{cases} 1 & (s > 0) \\ -1 & (s < 0) \end{cases}$$

$x_1$	$x_2$	$x_3$	$y$
0	0	0	1
0	1	1	-1
1	1	1	1
1	0	1	-1
0	0	1	1
0	1	0	-1



$$x_1 = w_0 + 0 \times w_1 + 0 \times w_2 + 0 \times w_3 = w_0 \quad \tau(w_0) = 1 \quad w_0 > 0$$

$$x_2 = w_0 + 0 \times w_1 + 1 \times w_2 + 1 \times w_3 \quad \tau(w_0 + w_2 + w_3) = -1$$

$$x_3 = w_0 + w_1 + w_2 + w_3 \quad \tau(w_0 + w_1 + w_2 + w_3) = 1$$

$$x_4 = w_0 + w_1 + w_3 \quad \tau(w_0 + w_1 + w_3) = -1$$

$$x_5 = w_0 + w_3 \quad \tau(w_0 + w_3) = 1$$

$$x_6 = w_0 + w_2 \quad \tau(w_0 + w_2) = -1 \quad w_2 < 0$$

$$\begin{aligned}
 & w_0 > 0 \quad w_2 < 0 \quad \frac{w_0 + w_3 > 0}{w_0 + w_3 + w_1 < 0} \\
 & w_1 < 0 \quad \Rightarrow 100\% \text{ 불가}
 \end{aligned}$$

$w_0 + w_3 + w_2 < 0$   
 $w_0 + w_3 + w_1 + w_2 > 0$

---

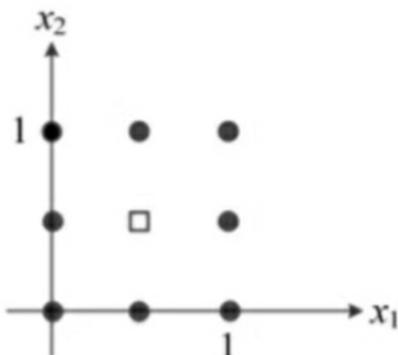
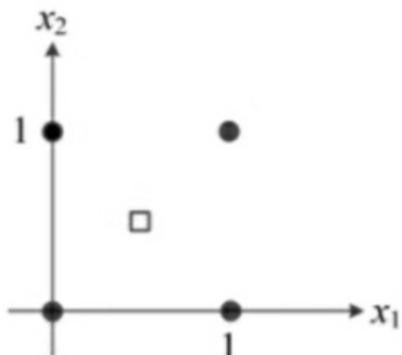
단증 퍼셉 트론시 최대 정확도  $\frac{5}{6}$

$$\text{ex) } w_0 = 0.5 \quad w_2 = -1 \quad w_3 = -0.4 \quad w_1 = -1$$

$$\begin{aligned}
 x_1 &= T(0.5) = 1, \quad x_2 = T(-1.1) = -1 \quad x_3 = T(-1.9) = -1 \\
 x_4 &= T(-0.9) = -1 \quad x_5 = T(0.1) = 1 \quad x_6 = T(-0.5) = -1
 \end{aligned}$$

$$\therefore y = 0.5 - x_1 - x_2 - 0.4x_3 \quad (\text{무수히 많은 퍼셉 트론 중 하나})$$

## 12 다음 분류 문제에 답하시오.

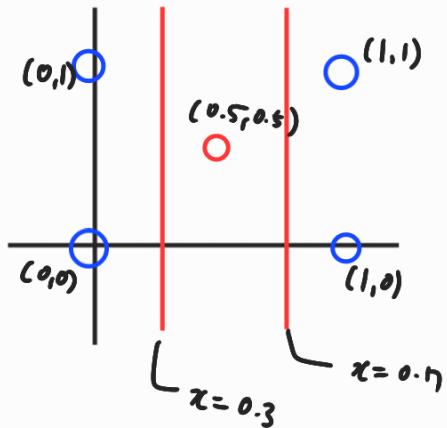


(1) 이 분류 문제를 해결하는 다층 퍼셉트론을 각각 제시하시오.

(2) 은닉 노드를 2개만 가진 퍼셉트론으로 오른쪽 상황을 해결할 수 있는지 답하시오. 가능하다면 은닉 노드 2개인 다층 퍼셉트론을 채택하는 것이 유리한지를 일반화 능력 측면에서 판단하시오.

**Hint** [그림 11-6]의 여백 개념을 사용하시오.

12 - 1 - 1



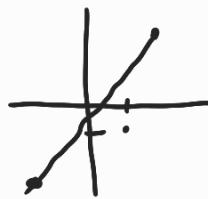
$$\begin{aligned} z_0 &= 1 - 0 \\ x_1 &\xrightarrow{w_1} z_1 \\ x_2 &\xrightarrow{w_2} z_1 \end{aligned}$$

$$\begin{aligned} x_1 w_1 + x_2 w_2 + w_0 &= z_1 \\ w_1 = 1, w_2 = 0, w_0 = -0.3 \text{ 일 때} \\ x_1 - 0.3 &= z_1 \end{aligned}$$

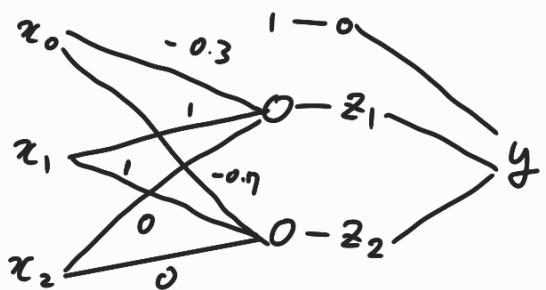
$x_1$	$x_2$	$z_1$	$t(z_1)$
0	0	-0.3	0
0	1	-0.3	0
1	0	0.7	1
1	1	0.7	1
0.5	0.5	0.2	1

$$\begin{aligned}
 & z_0 = 1 - 0 \\
 & z_1 \quad w_0 \quad w_1 \quad 0 - z_2 \\
 & z_2 \quad w_2 \quad 0 - z_2 \\
 & z_1 w_1 + z_2 w_2 + w_0 = z_2 \\
 & w_1 = 1 \quad w_2 = 0 \quad w_0 = -0.7 \text{ 이면} \\
 & z_1 - 0.7 = z_2
 \end{aligned}$$

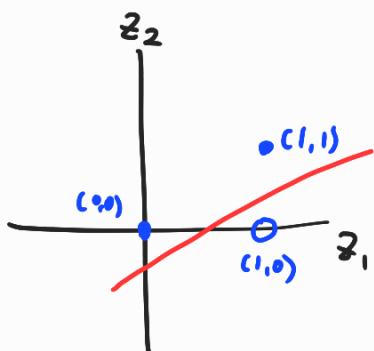
$x_1$	$x_2$	$z_2$	$T(z_2)$
0	0	-0.7	0
0	1	-0.7	0
1	0	0.3	1
1	1	0.3	1
0.5	0.5	-0.2	0



## 두 개의 직선 병렬 결합



$x_1$	$x_2$	$z_1$	$z_2$	$y$
0	0	0	0	0
1	0	0	0	0
0	1	1	1	1
1	1	1	1	1
0.5	0.5	1	0	0



$$w_0 + w_1 z_1 + w_2 z_2 = y$$

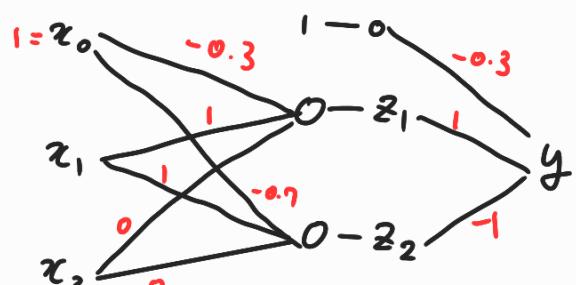
$$\text{ex)} \quad w_0 = -0.3 \quad w_1 = 1 \quad w_2 = -1$$

$$-0.3 + 0 + 0 = -0.3 \Rightarrow T(-0.3) = 0$$

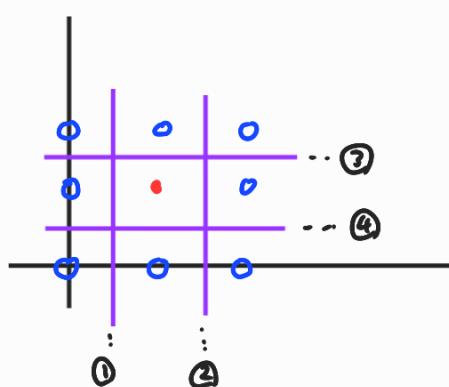
$$-0.3 + 1 = -0.3 \Rightarrow T(-0.3) = 0$$

$$-0.3 + 1 = 0.7 \Rightarrow T(0.7) = 1$$

$$\therefore y = -0.3 + z_1 - z_2 \quad (\text{무수히 많은 퍼셉트론 중 } \bar{z}_4)$$

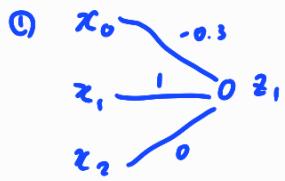


12-1 -2

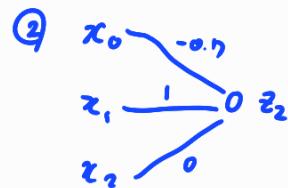


- ①  $x = 0.3$   
 ②  $x = 0.7$   
 ③  $y = 0.3$   
 ④  $y = 0.7$

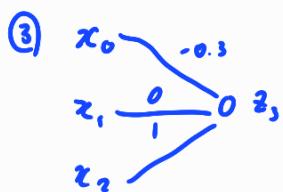
가정



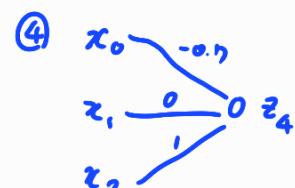
$x_1$	$x_2$	$z_1$	$T(z_1)$
0	0	-0.3	0
0	0.5	-0.3	0
0	1	-0.3	0
0.5	0	0.2	-
0.5	0.5	0.2	-
0.5	1	0.2	-
1	0	0.7	-
1	0.5	0.7	-
1	1	0.7	1



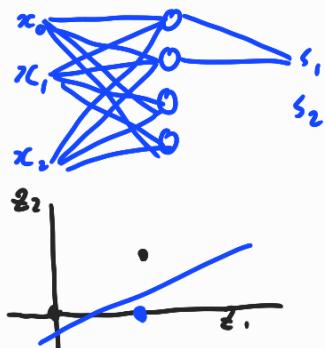
$x_1$	$x_2$	$z_2$	$T(z_2)$
0	0	-0.7	0
0	0.5	-0.7	0
0	1	-0.7	0
0.5	0	-0.7	0
0.5	0.5	-0.7	0
0.5	1	-0.7	0
1	0	0.3	-
1	0.5	0.3	-
1	1	0.3	1



$x_1$	$x_2$	$z_3$	$T(z_3)$
0	0	-0.3	0
0	0.5	0.2	-
0	1	0.7	-
0.5	0	-0.3	0
0.5	0.5	0.2	-
0.5	1	0.7	-
1	0	-0.3	0
1	0.5	0.2	-
1	1	0.7	1



$x_1$	$x_2$	$z_4$	$T(z_4)$
0	0	-0.7	0
0	0.5	-0.2	0
0	1	0.3	1
0.5	0	-0.7	0
0.5	0.5	-0.2	0
0.5	1	0.3	1
1	0	-0.7	0
1	0.5	-0.2	0
1	1	0.3	1



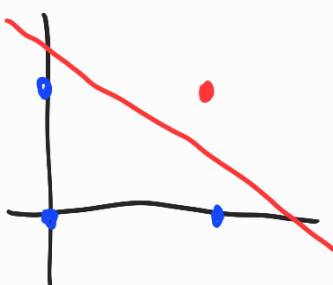
$x_1$	$x_2$	$z_1$	$z_2$	$z_3$	$z_4$	$s_1$	$s_2$	$T(s_1)$	$T(s_2)$	$b$	$T(b)$
0	0	0	0	0	0	-0.5	-0.5	0	0	1.5	1
0	0.5	0	0	1	0	-0.5	0.5	0	1	0.5	1
0	1	0	0	1	1	-0.5	-0.5	0	0	1.5	1
0.5	0	1	0	0	0	0.5	-0.5	1	0	0.5	1
0.5	0.5	1	0	1	0	0.5	0.5	1	1	-0.5	0
0.5	1	1	0	1	1	0.5	-0.5	1	0	0.5	1
1	0	0	1	0	0	-0.5	-0.5	0	0	1.5	1
1	0.5	0	1	0	0	-0.5	0.5	0	1	0.5	1
1	1	0	1	1	1	-0.5	0.5	0	0	1.5	1

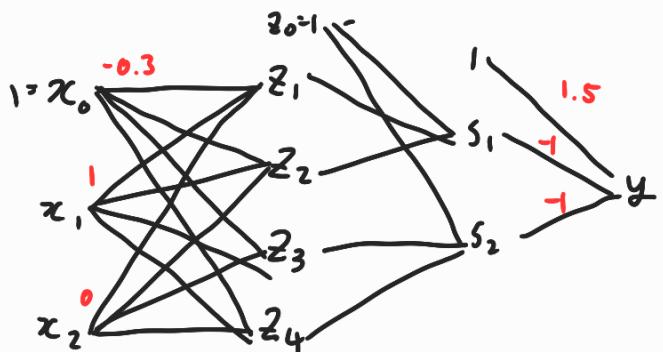
$$s_1: w_0 = -0.5, w_1 = 1 \quad w_2 = -1 \cdots \textcircled{1} \quad \textcircled{1} \quad \textcircled{2}$$

$$s_2: w_0 = -0.5, w_1 = 1 \quad w_2 = -1 \cdots \textcircled{2}$$

③

$$y: w_0 = 1.5 \quad w_1 = -1 \quad w_2 = -1 \cdots \textcircled{3}$$





input  $\rightarrow z_1$  :  $w_0 = -0.3, w_1 = 0, w_2 = 1$   
 input  $\rightarrow z_2$  :  $w_0 = -0.7, w_1 = 0, w_2 = 1$   
 input  $\rightarrow z_3$  :  $w_0 = -0.3, w_1 = 1, w_2 = 0$   
 input  $\rightarrow z_4$  :  $w_0 = -0.7, w_1 = 1, w_2 = 0$

$z_1, z_2 \rightarrow s_1$  :  $w_0 = -0.5 \quad w_1 = 1 \quad w_2 = -1$   
 $z_3, z_4 \rightarrow s_2$  :  $w_0 = -0.5 \quad w_1 = 1 \quad w_2 = -1$   
 $s_1, s_2 \rightarrow y$  :  $w_0 = 1.5 \quad w_1 = -1 \quad w_2 = -1$

12-2 12-1에서 보였듯이 원탁 노드 2개만 가진 노드로는 위의 분류 문제를 100% 분류 할 수 없다.

다면 hidden layer(은닉층)이 2개인 퍼셉트론으로는 분류가 가능하다.

퍼셉트론의 일 반화를 위해서는 매우 깊지 않은 노드로 퍼셉트론을 구성 해야지

Overfitting 문제가 발생하지 않는다.

SVM의 최소 마진 문제에서도 그렇듯 마진이 너무 train set에 최적화 되어 있으면 validation set이나 test set에서의 약간의 이상치에도 민감하게 반응하기 때문이다. 그렇다고 노드의 수를 너무 줄이면 underfitting issue가 발생한다.

따라서 적당한 노드의 수를 찾는 것이 최적이다.

13 [알고리즘 3-4]의 부연 설명에서 “X에서 임의로 샘플 하나를 뽑는” 랜덤 샘플링 버전을 소개하였다. 랜덤 샘플링 버전이 MNIST 훈련집합을 사용한다고 가정하자. MNIST는 숫자 부류 별로 6천 샘플을 가진다. [알고리즘 3-4]의 라인 2의 루프를 60만 번 반복한다고 가정한다. 10세대를 수행하는 셈이며 각 샘플은 평균 10번 선택된다. 다음 질문에 답하시오.

- (1) 어떤 샘플이 한 번도 선택되지 못할 확률은?
- (2) 어떤 부류에 속한 모든 샘플이 한 번도 선택되지 못할 확률은?
- (3) (1)과 (2)의 분석 결과를 토대로 랜덤 샘플링 방법의 유효성에 대한 의견을 제시하시오.

13-1) 샘플의 총 개수 60,000, 한번 시행시 샘플 0가 뽑히지 않을 확률  $(1 - \frac{1}{60000})$

$$\text{복원 추출 } 600,000 \text{ 번} \quad \therefore \left(1 - \frac{1}{60000}\right)^{600000} \approx 0.0000453$$

13-2) 각 부류의 샘플 개수 6000 개씩. 10개 (0~9까지)

$$\text{만약, } 0 \text{ 이 한번도 뽑히지 않을 확률} \Rightarrow \frac{9}{10}$$

$$60\text{만 번 복원 추출 } 0 \text{ 한 번도 뽑히지 않을 확률} \Rightarrow \left(\frac{9}{10}\right)^{600000} \approx 0$$

13-3) SGD의 랜덤 샘플링 방법은 유효하다.

13-1의 결과 샘플이 한번도 추출되지 않을 확률도 0에 가깝게 매우 작고

13-2의 결과를 보면 0에 가까워서 샘플의 편향성 문제는 발생하지 않을 것이다.