* Algorithm *

Gradient Scheme.

- A tool for finding the minimum of a function.

$$\text{grad } f(y) = \frac{df(y)}{dy} = \begin{bmatrix} \frac{\partial f}{\partial y_1} \\ \frac{\partial f}{\partial y_2} \\ \vdots \\ \frac{\partial f}{\partial y_n} \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ x_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}$$

- f : a scalar function of a vector argument.
- Each component of the gradient gives the rate of change of the function $\boxed{\text{in the direction of that component}}$.

* Perceptron.

- Adaptation { Reward. / Punishment.

⇓

More general approach?

<Family of iterative scheme>

- x_i ; the ith row of the $N \times (n+1)$ matrix X of the system of inequalities

$$XW > 0 \qquad X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{bmatrix}$$

$\uparrow$

C_1	C_2
$XW > 0$	$XW < 0$
$YW < 0$	$YW > 0$
+	−

⇒ If the patterns of C_2 are multiplied by -1, we obtain the equivalent condition

$$W^T X > 0$$

∘ $J(W,X) = (|W^TX| - W^TX)$

$\cdot\, |W^TX|$를 곱하고
$\cdot\, 4X^TX$로 나누고 $(X^TX > 0)$ $\Big\}$ ☆

* $$J(W,X) = \frac{1}{4X^TX}\left(|W^TX|^2 - |W^TX|\,W^TX\right)$$

$$\frac{\partial J(W,X)}{\partial W} = \frac{1}{4X^TX}\left[2|W^TX|X\,\mathrm{sgn}(W^TX) - |W^TX|X - (W^TX)X\,\mathrm{sgn}(W^TX)\right]$$

$$\mathrm{sgn}(W^TX) = \begin{cases} 1, & W^TX > 0 \\ -1, & W^TX \le 0 \end{cases}$$

$$= \begin{cases} \dfrac{1}{4X^TX}\left(2W^TX\cdot X - W^TX\cdot X - W^TXX\right) & (W^TX > 0) \\ \dfrac{1}{4X^TX}\left(2|W^TX|X(-1) - |W^TX|X + (W^TX)X\right) & (W^TX \le 0) \end{cases}$$

$$|\alpha| = \begin{cases} \alpha & \alpha > 0 \\ -\alpha & \alpha \le 0. \end{cases}$$

$$= \begin{cases} 0, & W^TX > 0 \\ \dfrac{W^TX}{X^TX}X, & W^TX \le 0 \end{cases}$$

$$= \frac{1}{2X^TX}\left[|W^TX|X\,\mathrm{sgn}(W^TX) - |W^TX|X\right]$$

∘ $$W(k+1) = W(k) - c\left\{\frac{\partial J(W,X)}{\partial W}\right\}_{W=W(k)}$$

$$= W(k) - c\left\{\frac{1}{2X^TX}\left[|W^TX|X\,\mathrm{sgn}(W^TX) - |W^TX|X\right]\right\}_{W=W(k)}$$

$$= W(k) - c\left\{\frac{1}{2X^T(k)X(k)}\left[|W^T(k)X(k)|X(k)\,\mathrm{sgn}(W^T(k)X(k)) - |W^T(k)X(k)|X(k)\right]\right.$$

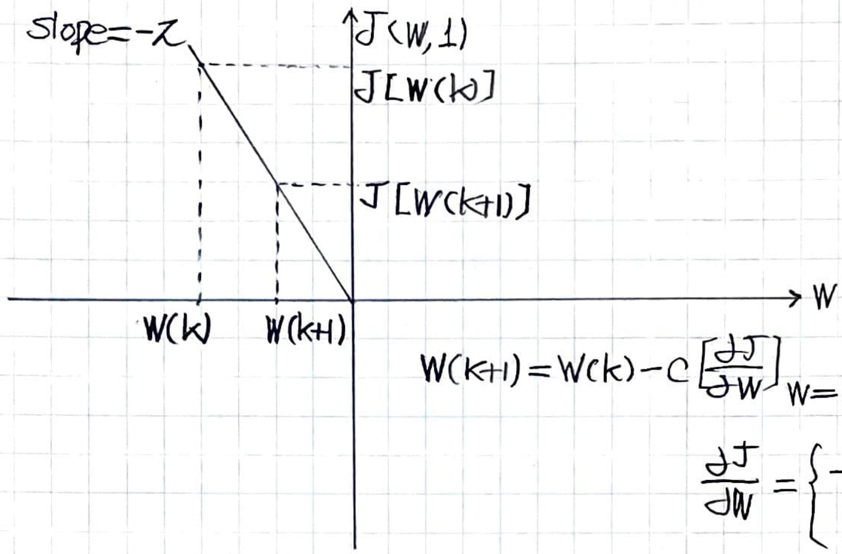
$$= W(k) + \frac{c}{2X^T(k)X(k)}\left[|W^T(k)X(k)|X(k) - |W^T(k)X(k)|X(k)\,\mathrm{sgn}(W^T(k)X(k))\right]$$

$$W(k+1) = W(k) + \frac{c\,|W^T(k)X(k)|}{2X^T(k)X(k)}\left[X(k) - X(k)\,\mathrm{sgn}\left(W^T(k)X(k)\right)\right]$$

$$= \begin{cases} W(k), & (W^T(k)X(k) > 0) \\ W(k) + \dfrac{c\,|W^T(k)X(k)|}{X^T(k)X(k)}X(k), & (W^T(k)X(k) \leq 0) \end{cases}$$

- This is a fractional-correlation algorithm.

Geometrical illustration of the gradient descent algorithm



$$W(k+1) = W(k) - c\left[\frac{\partial J}{\partial W}\right]_{W=W(k)}$$

$$\frac{\partial J}{\partial W} = \begin{cases} -z, & W \leq 0 \\ 0, & W > 0 \end{cases}$$

$$W(k+1) = \begin{cases} W(k) + zc & (W \leq 0) \\ W(k) & (W < 0) \end{cases}$$

- This scheme will eventually lead to a positive W, consequently, to the J_{min}.
- $J(W,X) = \left(|W^T X| - W^T X\right)\Big|_{X=1} \Rightarrow J(W,1) = |W^T| - W^T = \begin{cases} 0 & W > 0 \\ -2W^T & W \leq 0 \end{cases}$
- Clearly, there are as many curves as there are patterns in a problem.
- If the inequalities are consistent and a proper $J(W,X)$ is chosen, the algorithm $*$ will result in a solution.

$$* \quad W(k+1) = W(k) - c\left\{\frac{\partial J(W,X)}{\partial W}\right\}_{W=W(k)}$$

Course ______ Sheet ______ of ______

By ______ Date ______

$$J(W, X) = (|W^T X| - W^T X), \quad |W^T X| \geq 0$$

The minimum of this function is

$J(W, X) = 0$ and that this minimum resuls when $W^T X > 0$

The trivial case $W = 0$ is excluded.

- The approach employed below consists of incrementing W in the direction of the negative gradient of $J(W, X)$ in order to seek the minimum of the function.

- Let $W(k)$ represent the value W at the kth step, the general gradient descent algorithm becomes

$$W(k+1) = W(k) - c\left\{\frac{\partial J(W, X)}{\partial W}\right\}_{W = W(k)}$$

where $W(k+1)$ represents the new value of W, and $c > 0$ dictates the magnitude of the correction.

- It is noted that no corrections are made on W when $\left(\frac{\partial J}{\partial W}\right) = 0$, which is the condition for a minimum.

<Perceptron Algorithm> ; a reward-and-punishment iterative scheme.

- $J(W, X) = \frac{1}{2}(|W^T X| - W^T X)$

$$\frac{\partial J}{\partial W} = \frac{1}{2}[X \,\mathrm{sgn}(W^T X) - X], \quad \mathrm{sgn}(W^T X) = \begin{cases} 1, & \text{if } W^T X > 0 \\ -1, & \text{if } W^T X \leq 0 \end{cases}$$

$$= \begin{cases} 0 & (W^T X > 0) \\ -X & (W^T X \leq 0) \end{cases}$$

* Need to make a correction on the weight vector W whenever $W^T X \leq 0$,

(=′才 $W^T X < 0$로 정리 해유)

- $$W(k+1) = W(k) - c\left\{\frac{\partial J(W, X)}{\partial W}\right\}_{W = W(k)}$$

$$= W(k) - c \cdot \frac{1}{2}\left[X \,\mathrm{sgn}(W^T X) - X\right]_{W = W(k)}$$

$$= W(k) - \frac{c}{2}\left[X(k) \,\mathrm{sgn}(W(k)^T X(k)) - X(k)\right]$$

$$= W(k) + \frac{c}{2}\left[X(k) - X(k) \,\mathrm{sgn}(W(k)^T X(k))\right]$$

$$= \boxed{\begin{cases} W(k), & (W^T X > 0) \\ W(k) + cX(k), & (W^T X \leq 0) \end{cases}}$$

($c > 0$ and $W(1)$ is arbitrary)