

#### Question. 4-06

Question. 4-04와 Question. 4-05의 결과를 토대로  $|x| \geq 1$ 인 경우와  $0 \leq |x| < 1$ 인 경우에서  $\theta_1$ 과  $\theta_0$  중 어떤 learnable parameter의 학습이 더 주도적으로 일어나는지 설명하시오.

또한, 이를 토대로  $|x| \geq 1$ 인 경우와  $0 \leq |x| < 1$ 인 경우에서  $\theta_1, \theta_0$  중 어떤 것이 더 발산할 위험성이 높은지 설명하시오.

1)  $|x| \geq 1$ 인 경우  $\theta_1$ 의 update값인  $2xx(y - \theta_1 x - \theta_0)$ 가 항상  $\theta_0$ 의 update값인  $2x(y - \theta_1 x - \theta_0)$ 보다 크므로  $\theta_1$ 이  $|x|$ 에 비례하여 dominant하게 학습이 이루어진다.  
따라서  $x$ 가 자칫 커질 경우  $\theta_1$ 이 발산할 위험성이 커진다.

$0 \leq |x| < 1$ 인 경우  $\theta_1$ 의 update값인  $2xx(y - \theta_1 x - \theta_0)$ 가 항상  $\theta_0$ 의 update값인  $2x(y - \theta_1 x - \theta_0)$ 보다 작으므로  $\theta_0$ 가  $|x|$ 에 비례하여 dominant하게 학습이 이루어진다.  
따라서 이 경우엔 반대로  $\theta_0$ 가 발산할 위험성이 커진다.