1. Gradient Descent.

(1) $a=2$

$l_a(w) = 2w^2$,

$\therefore \dfrac{d\, l_a(w)}{dw} = 4w$

for: $w = w'$, next $w'' = w' - \eta 4 w'$

① for $\eta = \frac{1}{2}$,

$w_0 = 1$,	$l = 2$	the value of
$w_1 = 1 - \frac{1}{2} \times 4 \times 1 = -1$	$l = 2$	
$w_2 = -1 - \frac{1}{2} \times 4 \times (-1) = 1$	$l = 2$	w is $[1, -1, 1, -1$
$w_3 = 1 - \frac{1}{2} \times 4 \times 1 = -1$	$l = 2$	
$w_4 = -1 - \frac{1}{2} \times 4 \times (-1) = 1$	$l = 2$	$\ldots w_{2k+1} = -1$,
$\ldots$	$\ldots$	$w_{2k} = 1]$.
$w_{2k+1} = -1 \quad (k \geq 0, k \in N^+)$.	$l = 2$	
$w_{2k} = 1$	$l = 2$.	the value of

$l_a(w) = 2$ which is fixed.

② for $\eta = 2$

$w_0 = 1$ $\qquad L = 2$

$w_1 = 1 - 2 \times 4 \times 1 = -7$ $\qquad L = 98$

$w_2 = -7 - 2 \times 4 \times (-7) = 49$ $\qquad L = 4802$

$w_3 = 49 - 2 \times 4 \times (49) = -343$ $\qquad L = 235298$.

......

$w_n = (-7)^n$ $\qquad L = 2 \times 49^n$.

the value of w is $[(-7)^0, (-7)^1, (-7)^2, \cdots, (-7)^n]$

the value of $L_a(w)$ is 2×49^n. $(w_i, i \in [0, n])$

Conclusion: from ①, $\eta = \frac{1}{2}$, the loss function doesn't decrease; from ② $\eta = 2$, the loss function is bigger and bigger. the learning rates of them are two large, so they don't get solutions.

(2). $L_d(w) = aw^2 \Rightarrow \dfrac{d\,L_d(w)}{w} = 2aw$

$\therefore \dfrac{2aw_1}{2aw_0} \leq -1$

$\therefore \dfrac{w_0 - 2a\eta w_0}{w_0} \leq -1$

$\therefore \eta \geq \dfrac{1}{a}$

(3). $\dfrac{d\,L_d(w)}{w} = 2aw$

$i=1, w = w_0$: $w_1 = w_0 - 2a\eta w_0 = w_0(1-2a\eta)$

$i=2, w = w_1$: $w_2 = w_1 - 2a\eta w_1 = w_1(1-2a\eta)$

$= w_0(1-2a\eta)^2$

$n=i$, $w=w_{i-1}$: $w_i = w_0(1-2a\eta)^i$

$\therefore \quad w_i - w^* = w_0(1-2a\eta)^i - w^*$

又$\because \quad |w_i - w^*| < \varepsilon$

$\therefore \quad -\varepsilon < w_i - w^* < \varepsilon$

$\therefore$ 解 $-\varepsilon < w_0(1-2a\eta)^i - w^* < \varepsilon$

$\therefore$

$$\log_{(1-2a\eta)} \frac{w^*-\varepsilon}{w_0} < i < \log_{1-2a\eta} \frac{\varepsilon + w^*}{w_0}$$