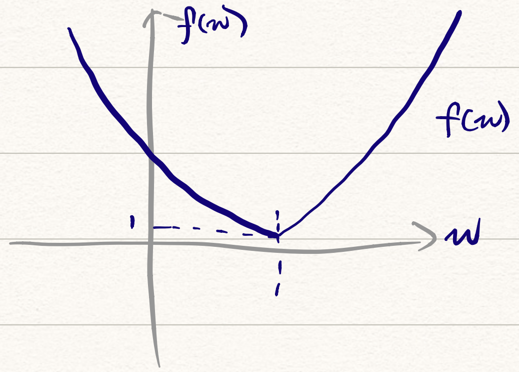
①



(a) $f(w)$ is convex. Any tangent line of $f(w)$ lies entirely under $f(w)$, since $f(w)$ always concave up, except for $w=1$.

(b) No. $f(w)$ is not differentiable at $w=1$, because there is a corner.

(c)

$-2.$

$$\frac{d}{dw} f(w) = \begin{cases} -2e^{-2(w-2)}, & w<1 \\ e^{w-1}, & w>1 \\ [-2e^{-2(-1)}, e^{1-1}] = [-2e^{2}, 1], & w=0 \end{cases}$$

②

(a) define subgradient $$v_i(w) = \begin{cases} -b_i x_i, & b_i x_i^T w<1 \\ 0, & b_i x_i^T w \geq 1 \end{cases} = -d_i x_i I\{d_i x_i^T w<1\}$$

cost function $$f(w) = \sum_{i=1}^{m} (1-b_i x_i^T w)_+$$

$$\nabla f(w)|_{w^{(k)}} = \sum_{i=1}^{m} (-b_i x_i I\{b_i x_i^T \tau^T w^{(k)}<1\})$$

Gradient descent $w^{(k+1)} = w^{(k)} - \tau \nabla f(w)|_{w^{(k)}}$, where τ is step size

(b) Gradient descent has arrived at or close to global min.

③ $w^{(k+1)} = w^{(k)} + \tau (d_{ik} - x_{ik}^T w^{(k)}) x_{ik} - \frac{\tau \varepsilon}{2N} \operatorname{sign}(w^{(k)}).$

$w^1 = w^0 + 1 \cdot (1 - [1 \ -1] \, 0] \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \frac{2 \cdot 1}{2 \cdot 4} \cdot 1 \quad = \begin{bmatrix} 3/4 \\ -5/4 \end{bmatrix}$

$w^2 = \begin{bmatrix} 3/4 \\ -5/4 \end{bmatrix} + 1 \cdot (2 - [1 \ -2] \begin{bmatrix} 3/4 \\ -5/4 \end{bmatrix}) \begin{bmatrix} 1 \\ -2 \end{bmatrix} - \frac{2 \cdot 1}{2 \cdot 4} \cdot 1$

$= \begin{bmatrix} \frac{-7}{12} \\ \frac{-5}{12} \end{bmatrix}$