475 Machine Learning HW1

Yuanwen Liu

04/11/2020

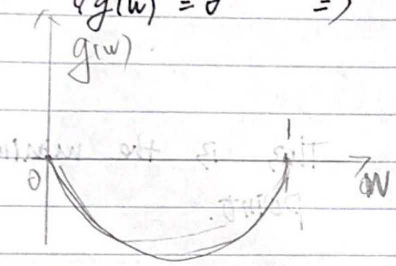
3.1 3.3 3.5 3.6 3.8

3.1 (a) $g(w) = w\log(w) + (1-w)\log(1-w)$

$\nabla g(w) = \log(w) + w \cdot w/w \;,\; -\log(1-w) + (1-w)/[-(1-w)]$

$= \log(w) - \log(1-w)$

$\nabla g(w) = 0 \Rightarrow w = 1-w \Rightarrow w = 0.5$



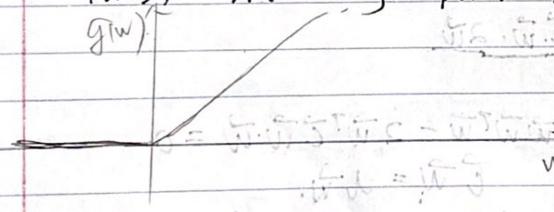
$w = 0.5$ is minimum point.

(b) $g(w) = \log(1+e^w)$

$\nabla g(w) = \frac{e^w}{(1+e^w)} = \frac{1}{(\frac{1}{e^w}+1)}$

$w \to -\infty \quad e^w \to 0 \;, \quad \frac{1}{e^w} \to \infty \Rightarrow \nabla g(w) \to 0$

Thus, stationary point is at minus infinite.



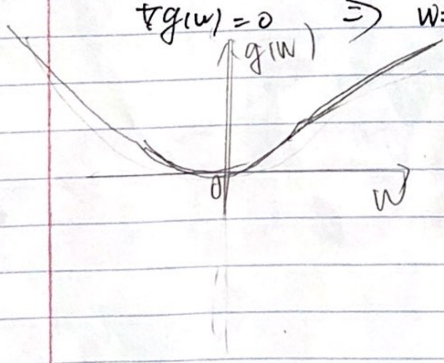
Thus, it's the minimum point.

(c) $g(w) = w\tanh(w)$

$\nabla g(w) = \tanh(w) + w \cdot [1-\tanh^2(w)] = -i\tan(iw) + w\cdot[1+\tan^2(iw)]$

$= \frac{e^w - e^{-w}}{e^w + e^{-w}} + w \cdot \frac{4}{(e^w+e^{-w})^2} = \frac{e^{2w} - e^{-2w} + 4w}{(e^w+e^{-w})^2}$

$\nabla g(w) = 0 \Rightarrow w = 0$ This is the stationary point.



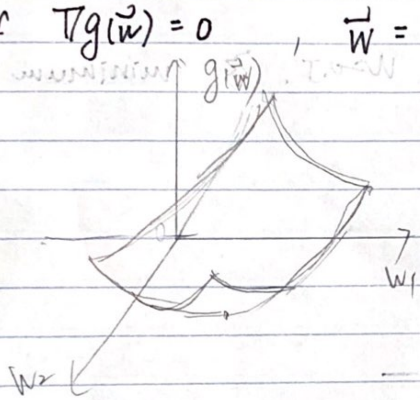
This is the minimum point.

(d) $g(\vec{w}) = \frac{1}{2}\vec{w}^T\vec{C}\vec{w} + \vec{b}^T\vec{w}$. where $\vec{C} = \begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}$ $\vec{b} = \begin{bmatrix}1\\ 1\end{bmatrix}$

$\nabla g(\vec{w}) = \vec{C}\vec{w} + \vec{b}$

$= \begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}\vec{w} + \begin{bmatrix}1\\ 1\end{bmatrix} = \begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}\begin{bmatrix}w_1\\ w_2\end{bmatrix} + \begin{bmatrix}1\\ 1\end{bmatrix} = \begin{bmatrix}2w_1 + w_2\\ w_1 + 3w_2\end{bmatrix} + \begin{bmatrix}1\\ 1\end{bmatrix}$

If $\nabla g(\vec{w}) = 0$, $\vec{w} = \begin{bmatrix}-\frac{2}{5}\\ -\frac{1}{5}\end{bmatrix}$



This is the minimum point.

3.3 $g(\vec{w}) = \frac{\vec{w}^T\vec{C}\vec{w}}{\vec{w}^T\vec{w}}$

$\nabla g(\vec{w}) = \frac{2\vec{C}\vec{w}\cdot\vec{w}^T\vec{w} - \vec{w}^T\vec{C}\vec{w}\cdot 2\vec{w}}{\vec{w}^T\vec{w}\cdot\vec{w}^T\vec{w}}$

Let $\nabla g(\vec{w}) = 0$ $\quad 2\vec{C}\vec{w}\vec{w}^T\vec{w} - 2\vec{w}^T\vec{C}\vec{w}\cdot\vec{w} = 0$

$\vec{C}\vec{w} = \lambda\vec{w}$.

Stationary point is the eigen-vector of $\vec{C}$.

3.5 $\nabla g(w) = \frac{1}{50}(4w^3 + 2w + 10)$