Machine Learning Worksheet 3

Tomas Ladek, Michael Kratzer
3602673,
tom.ladek@tum.de,

Problem 1

$$\begin{aligned}\frac{\partial}{\partial\theta}\theta^t(1-\theta)^h &= t\theta^{t-1}(1-\theta)^h - \theta^t h(1-\theta)^{h-1}\\ &= \theta^{t-1}(1-\theta)^{h-1}\big(t(1-\theta)-h\theta\big)\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}\theta^t(1-\theta)^h &= \frac{\partial}{\partial\theta}\theta^{t-1}(1-\theta)^{h-1}\big(t(1-\theta)-h\theta\big)\\ &= -2\,t\,h\,\theta^{t-1}(1-\theta)^{h-1} + t(t-1)\theta^{t-2}(1-\theta)^h + h(h-1)\theta^t(1-\theta)^{h-2}\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial\theta}\log\theta^t(1-\theta)^h &= \frac{\partial}{\partial\theta}\big(t\log\theta + h\log(1-\theta)\big)\\ &= \frac{t}{\theta} - \frac{h}{1-\theta}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}\log\theta^t(1-\theta)^h &= \frac{\partial}{\partial\theta}(\frac{t}{\theta} - \frac{h}{1-\theta})\\ &= \frac{h}{(1-\theta)^2} - \frac{t}{\theta^2}\end{aligned}$$

Problem 2

The local extremum of $\log f(\theta)$ can determined by setting the first derivative to zero, i.e.

$$\frac{d}{d\theta}\log\big(f(\theta)\big) \stackrel{!}{=} 0$$

Evaluating this with the logarithm derivative rule leads to

$$\frac{\frac{d}{d\theta}f(\theta)}{f(\theta)} = 0 \iff \frac{d}{d\theta}f(\theta) = 0$$

Which is also the formula for the extremum of a differentiable positive function $f(\theta)$. That means that taking the logarithm of a function preserves its extremum points (in particular its local maximum points). Considering the results from the first exercise, it can be easier to differentiate the logarithm of a function instead of the function itself, when one is interested only in the location of the maximum and not its value.

Problem 3