Machine Learning: Homework #6

Due on December 04, 2017 at 09:59am

Professor Dr. Stephan Guennemann

Marc Meissner - 03691673

Problem 1

Prove or disprove whether the following functions are convex on the given set D:

i) $f(x, y, z) = 3x + e^{y+z} - \min(-x^2, \log(y))$ and $D = (-100, 100) \times (1, 50) \times (10, 20)$
ii) $f(x, y) = yx^3 - 2yx^2 + y + 4$ and $D = (-10, 10) \times (-10, 10)$
iii) $f(x) = log(x) + x^3$ and $D = (1, \infty)$
iv) $f(x) = -\min(2\log(2x), -x^2 + 4x - 32)$ and $D = \mathbb{R}^+$

Solution
i)
One can immediately see that $\min(-x^2, \log(y)) = -x^2$ for the given set D. Ergo:
$f(x, y, z) = 3x + e^{y+z} - \min(-x^2, \log(y)) = 3x + e^{y+z} + x^2$
Since these are just additions of convex functions, f is convex.

ii)
Let $\vec{x_1} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\vec{x_2} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\lambda = \frac{1}{2}$. Then

$\lambda f(\vec{x_1}) + (1 - \lambda) f(\vec{x_2}) \geq f(\lambda \vec{x_1} + (1 - \lambda)\vec{x_2})$

would have to hold for the function to be convex. But

$\lambda f(\vec{x_1}) + (1 - \lambda) f(\vec{x_2}) = 3$ and $f(\lambda \vec{x_1} + (1 - \lambda)\vec{x_2}) = 4.1875$

Here, $\lambda f(\vec{x_1}) + (1 - \lambda) f(\vec{x_2}) < f(\lambda \vec{x_1} + (1 - \lambda)\vec{x_2})$ and thus the function is not convex.

iii)
If the second derivative is non-negative for the given set D, it is convex. Calculating the derivatives and checking the second one proves the convexity of f:

$\frac{df}{dx} = \frac{1}{x} + 3x^2$ and $\frac{d^2f}{dx^2} = -\frac{1}{x^2} + 6x > 0$ for $x \in [1; \infty]$

iv)
Intuitively, the log function goes to $-\infty$ for small x values:

$lim_{x \to 0} - \min(2\log(2x), -x^2 + 4x - 32) = -\min(-\infty, -32) = \infty$

The parabola is convex anyway. At one point, the log function becomes smaller than the parabola. Since the negative log is convex for $0 < x < 1$, f is convex as well.

Problem 2

Prove the following statement: Let $f_1 : \mathbb{R}^d \to \mathbb{R}$ and $f_2 : \mathbb{R}^d \to \mathbb{R}$ be convex functions, then $h(x) = f_1(x) + f_2(x)$ is also a convex function.

Solution
Using the definition of convexity:

$\lambda h(x) + (1-\lambda)h(y)$
$= \lambda(f_1(x) + f_2(x)) + (1-\lambda)(f_1(y) + f_2(y))$
$= \lambda f_1(x) + (1-\lambda)f_1(y) + \lambda f_2(x) + (1-\lambda)f_2(y)$
$\geq f_1(\lambda x + (1-\lambda)y) + f_2(\lambda x + (1-\lambda)y)$
$= h(\lambda x + (1-\lambda)y)$

q.e.d. bish

Problem 3

Given two convex functions $f_1 : \mathbb{R} \to \mathbb{R}$ and $f_2 : \mathbb{R} \to \mathbb{R}$, prove or disprove that the function $h(x) = f_1(x) \cdot f_2(x)$ is also convex.

Solution
Simple counterexample: imagine two linear (and thus convex) functions $f_1(x) = x$ and $f_2(x) = -x$. The product $g(x) = x \cdot (-x) = -x^2$ is obviously not convex.

Problem 4

Prove that for convex functions each local minimum is a global minimum. More specifically, given a convex function $f : \mathbb{R}^N \to \mathbb{R}$, prove that if $\nabla f(\theta^*) = 0$ then θ^* is a global minimum.

Solution

Problem 5

Load the notebook 06*_hw_optimization_logistic_regression.ipynb* from Piazza. Fill in the missing code and run the notebook. Convert the evaluated notebook to pdf and add it to the printout of your homework.

Problem 5 continued on next page. . .

Solution