Machine Learning Exercise Sheet 06

Optimization

Exercise sheets consist of two parts: homework and in-class exercises. You solve the homework exercises on your own or with your registered group and upload it to Moodle for a possible grade bonus. The in-class exercises will be solved and explained during the tutorial. You do not have to upload any solutions of the in-class exercises.

In-class Exercises

Problem 1: Prove or disprove whether the following functions $f : D \to \mathbb{R}$ are convex

a) $D = (1, \infty)$ and $f(x) = \log(x) - x^3$,

b) $D = \mathbb{R}^+$ and $f(x) = -\min(\log(3x + 1), -x^4 - 3x^2 + 8x - 42)$,

c) $D = (-10, 10) \times (-10, 10)$ and $f(x, y) = y \cdot x^3 - y \cdot x^2 + y^2 + y + 4$.

Problem 2: Prove that the following function (the loss function of logistic regression) $f : \mathbb{R}^d \to \mathbb{R}$ is convex:

$$f(\boldsymbol{w}) = -\ln p\big(\boldsymbol{y} \mid \boldsymbol{w}, \boldsymbol{X}\big) = -\sum_{i=1}^{N} \big(y_i \ln \sigma(\boldsymbol{w}^T \boldsymbol{x}_i) + (1 - y_i) \ln(1 - \sigma(\boldsymbol{w}^T \boldsymbol{x}_i))\big) .$$

Problem 3: Prove that for differentiable convex functions each local minimum is a global minimum. More specifically, given a differentiable convex function $f : \mathbb{R}^d \to \mathbb{R}$, prove that

a) if $\boldsymbol{x}^*$ is a local minimum, then $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$.

b) if $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$, then $\boldsymbol{x}^*$ is a global minimum.

Upload a single PDF file with your homework solution to Moodle by 16.12.2020, 23:59 CET. We recommend to typeset your solution (using LaTeX or Word), but handwritten solutions are also accepted. If your handwritten solution is illegible, it won't be graded and you waive your right to dispute that.

Homework

1 Convexity of functions

Problem 4: Assume that $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are convex functions. Prove or disprove the following statements:

a) The function $h(x) = g(f(x))$ is convex.

b) The function $h(x) = g(f(x))$ is convex if g is non-decreasing.

Note: For this exercise you are not allowed to use the convexity preserving operations from the lecture.

2 Optimization / Gradient descent

Problem 5: You are given the following objective function $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x_1, x_2) = 0.5x_1^2 + x_2^2 + 2x_1 + x_2 + \cos(\sin(\sqrt{\pi})).$$

a) Compute the minimizer $\boldsymbol{x}^*$ of f analytically.

b) Perform 2 steps of gradient descent on f starting from the point $\boldsymbol{x}^{(0)} = (0, 0)$ with a constant learning rate $\tau = 1$.

c) Will the gradient descent procedure from Problem b) ever converge to the true minimizer $\boldsymbol{x}^*$? Why or why not? If the answer is no, how can we fix it?

Problem 6: Load the notebook `exercise_06_notebook.ipynb` from Piazza. Fill in the missing code and run the notebook. Export (download) the evaluated notebook as PDF and add it to your submission.

Note: We suggest that you use Anaconda for installing Python and Jupyter, as well as for managing packages. We recommend that you use Python 3.

For more information on Jupyter notebooks, consult the Jupyter documentation. Instructions for converting the Jupyter notebooks to PDF are provided on Piazza.

Problem 7: Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the following convex function:

$$f(x_1, x_2) = e^{x_1+x_2} - 5 \cdot \log(x_2)$$

a) Consider the following shaded region $S \subset \mathbb{R}^2$. Is this region convex? Why?

b) Assume that we are given an algorithm $\texttt{ConvOpt}(f, D)$ that takes as input a convex function f and convex region D, and returns the minimum of f over D. Using the `ConvOpt` algorithm, how would you find the global minimum of f over the shaded region S?