Convex Optimization for Computer Vision
Lecture: M. Möller
Exercises: J. Geiping
Summer Semester 2019

Universität Siegen
Department ETI
Computer Vision

Weekly Exercises 6

Room: H-A 7116
Thursday, 13.06.2019, 8:30-10:00,
Submission deadline: Wednesday, 12.06.2019, 18:00

Theory

Exercise 1 (2 points)**.** Let $G : \mathbb{R}^n \to \mathbb{R}^m$ and $F : \mathbb{R}^m \to \mathbb{R}^r$. Show that if one of the operators is a contraction and the other one is non-expansive, then $(F \circ G)$ is a contraction, too.

Exercise 2 (4 points)**.** A classical task in machine learning is *multinomial logistic regression*. Assume you are given training data consisting of n pairs $X_{i,:} \in \mathbb{R}^m$ and a corresponding label $t_i \in \{1, \dots, c\}$, and want to find a mapping that predicts the label based on the input data $X_{i,:}$. The idea of the multinomial logistic regression method is to look for some *weights* $W \in \mathbb{R}^{m,c}$ and *biases* $b \in \mathbb{R}^c$ such that the function

$$\mathcal{N}(x; W, b) = (xW)^T + b \in \mathbb{R}^c$$

maps feature vectors $x \in \mathbb{R}^m$ to the *scores* for each class in $\{1, \dots, c\}$, which correspond to the probabilities of x being in the respective class. Now we want to optimize W and b by minimizing the following energy function

$$E(W, b) = \frac{1}{n} \sum_{i=1}^{n} \ell(W, b, X_{i,:}, t_i) + \frac{\lambda}{2} \|W\|_F^2 + \frac{\lambda}{2} \|b\|_2^2. \tag{1}$$

The loss function ℓ is given by

$$\ell(W, b, x, t) = -\log \left(\frac{\exp(xW_{:,t} + b_t)}{\sum_{j=1}^{c} \exp(xW_{:,j} + b_j)} \right) \tag{2}$$

Determine the gradient of E with respect to W and b! (Since this gradient is crucial for the programming exercise, please contact me if you get stuck!)

Programming: Multinomial logistic regression

Exercise 3 (8 Points)**.** Your task is to minimize the energy given in (1) on an example classification problem, namely classifying wines by winery based on its chemical characteristics. As explained at `https://de.mathworks.com/help/nnet/examples/`

`wine-classification.html` the MATLAB *wine_ dataset* contains 178 examples of wines from three different wineries that have been classified with respect to 13 different attributes. Your task is to find weights $W \in \mathbb{W}^{13\times 3}$ and biases $b \in \mathbb{W}^{1\times 3}$ that minimize (1). We will hold back 20% of the overall data to test up to which accuracy you can assign new wines to the respective wineries they came from.

Use `logRegQuickTemplate.m` as starting point for your implementation.