Machine Learning Homework Sheet 06

Constrained Optimization and SVM

1 Constrained Optimization

Problem 1: Solve the following constrained optimization problem using the recipe described in the lecture (slide 17).

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{\theta}) = \theta_1 - \sqrt{3}\theta_2 \\ \text{subject to} \quad & f_1(\boldsymbol{\theta}) = \theta_1^2 + \theta_2^2 - 4 \leq 0 \end{aligned}$$

2 Projected Gradient Descent

Problem 2: Given is the following (convex) domain defined by a set of linear constraints

$$\mathcal{X} \subset \mathbb{R}^2 = \{\boldsymbol{x} \in \mathbb{R}^2 : (x_1 + x_2 \leq 4) \wedge (0 \leq x_1 \leq 3) \wedge (0 \leq x_2 \leq 2.5)\}.$$

a) Visualize the set $\mathcal{X}$.

b) Derive a closed form for the projection $\pi_{\mathcal{X}}(\boldsymbol{p}) = \arg\min_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x} - \boldsymbol{p}||_2^2$. That is, given an arbitrary point $\boldsymbol{p} \in \mathbb{R}^2$, what is its projection on $\mathcal{X}$?

   Hint: For one part of $\mathbb{R}^2$ you might want to use the line projection $\pi_{\text{line}}(\boldsymbol{p}) = \boldsymbol{a} + \frac{(\boldsymbol{p}-\boldsymbol{a})^T(\boldsymbol{b}-\boldsymbol{a})}{\|\boldsymbol{b}-\boldsymbol{a}\|_2^2}(\boldsymbol{b}-\boldsymbol{a})$, where $\boldsymbol{a} \in \mathbb{R}^2$ and $\boldsymbol{b} \in \mathbb{R}^2$ specify the line.

c) Given is the following constrained optimization problem:

$$\min_{\boldsymbol{x}} (x_1 - 2)^2 + (2x_2 - 7)^2,$$
$$\text{subject to } \boldsymbol{x} \in \mathcal{X}.$$

   Perform two steps of projected gradient descent starting from the point $\boldsymbol{x}^{(0)} = (2.5, 1)^T$. Use a constant learning rate/step size of $\tau = 0.05$.

3 SVM

Problem 3: Explain the similarities and differences between the SVM and perceptron algorithms.

Problem 4: Show that the duality gap is zero for SVM.

Upload a single PDF file with your solution to Moodle by 2.12.2018, 23:59pm 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.

Problem 5: Recall that the dual function for SVM (slide 41) can be written as

$$g(\boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{\alpha}^T \boldsymbol{Q} \boldsymbol{\alpha} + \boldsymbol{\alpha}^T \mathbf{1}_N$$

(a) Show how the matrix $\boldsymbol{Q}$ can be computed. *(Hint: You might want to use Hadamard product, denoted as $\odot$).*

(b) Prove that the matrix $\boldsymbol{Q}$ is negative (semi-)definite.

(c) Explain what the negative (semi-)definiteness means for our optimization problem. Why is this property important?

Problem 6: Download the notebook `homework_06_notebook.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 (see printing instructions inside the notebook).