Machine Learning Exercise Sheet 07

Constrained Optimization

Homework

1 Constrained Optimization / Projected GD

Problem 1: Given is the following domain $\mathcal{X} \subset \mathbb{R}^2$ defined by a set of linear constraints

$$\begin{aligned}\mathcal{X} = \{\boldsymbol{\theta} \in \mathbb{R}^2 :\ & \theta_1 + \theta_2 \leq 12,\\ & -\theta_1 + 2\theta_2 \geq -3,\\ & -5\theta_1 + 3\theta_2 \leq -4,\\ & \theta_1 \geq 2,\ \theta_2 \geq 2\}\end{aligned}$$

and $f_0 : \mathbb{R}^2 \to \mathbb{R}$ with $f(\boldsymbol{\theta}) = 2\theta_1 - 3\theta_2$.

Visualize set $\mathcal{X}$ and find the minimizer $\boldsymbol{\theta}_{\text{min}} \in \mathbb{R}^2$ and maximizer $\boldsymbol{\theta}_{\text{max}} \in \mathbb{R}^2$ of f_0 on $\mathcal{X}$ as well as it's minimum and maximum values. There is no need for a rigorous derivation in this exercise, but your plot should be nice and understandable and you still have to provide an explanation of your steps as well as clearly mention what results you apply.

Problem 2: Given is the following domain $\mathcal{X} \subset \mathbb{R}^2$ defined by a set of linear constraints

$$\begin{aligned}\mathcal{X} = \{\boldsymbol{\theta} \in \mathbb{R}^2 :\ & \theta_1 + \theta_2 \leq 4,\\ & 0 \leq \theta_1 \leq 3,\ 0 \leq \theta_2 \leq 2.5\}.\end{aligned}$$

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

Hint: Use your plot from a) to divide $\mathbb{R}^2$ into regions. For one part of $\mathbb{R}^2$ you might want to use the following formula for projection on a hyperplane $\mathcal{X}_{a,b} = \{\boldsymbol{\theta} \in \mathbb{R}^2 : \boldsymbol{a}^T\boldsymbol{\theta} + \boldsymbol{b} = 0\}$ defined by vectors $\boldsymbol{a} \in \mathbb{R}^2$ and $\boldsymbol{b} \in \mathbb{R}^2$:

$$\pi_{\mathcal{X}_{a,b}}(\boldsymbol{p}) = \boldsymbol{a} + \frac{(\boldsymbol{p} - \boldsymbol{a})^T(\boldsymbol{b} - \boldsymbol{a})}{||\boldsymbol{b} - \boldsymbol{a}||_2^2}(\boldsymbol{b} - \boldsymbol{a})$$

c) Perform two steps of projected gradient descent starting from the point $\boldsymbol{\theta}^{(0)} = (2.5, 1)^T$ with a constant learning rate of $\tau = 0.05$ for the following constrained optimization problem:

$$\begin{aligned}\text{minimize}_{\boldsymbol{\theta}} \quad & (\theta_1 - 2)^2 + (2\theta_2 - 7)^2\\ \text{subject to} \quad & \boldsymbol{\theta} \in \mathcal{X}\end{aligned}$$

Upload a single PDF file with your homework solution to Moodle by 01.12.2019, 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.

2 Lagrangian / Duality

Problem 3: Solve the following constrained optimization problem in $\mathbb{R}^2$ using the recipe described in the lecture (Slide 24).

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

Problem 4: Given N data points $\boldsymbol{x}_i \in \mathbb{R}^d$ and their labels $y_i \in \{-1, 1\}$. Use results you know from the lecture and show that strong duality holds for the following problem.

$$\begin{aligned} \text{minimize}_{\boldsymbol{w}, b} \quad & \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} \\ \text{subject to} \quad & y_i(\boldsymbol{w}^T\boldsymbol{x}_i + b) - 1 \geq 0 \quad \text{for } i = 1, \dots, N \end{aligned}$$

Note: by solving this problem we can find the separating hyperplane between points from different classes with the maximum margin, we will talk about it in the lecture when we cover SVMs.

In-class Exercises

3 Duality

Problem 5: Given N numbers $x_1, ..., x_N \in \mathbb{R}$, construct a minimization problem such that its optimal value is $\max(x_1, \ldots, x_N)$ and derive the Lagrange dual problem.

Problem 6: Given N numbers $x_1, ..., x_N \in \mathbb{R}$, construct a maximization problem such that its optimal value is the sum of the k largest values and derive the Lagrange dual problem.

Problem 7: Given $\boldsymbol{c} \in \mathbb{R}^d$, $\boldsymbol{b} \in \mathbb{R}^M$ and $\boldsymbol{A} \in \mathbb{R}^{M \times d}$, derive the Lagrange dual problem of

$$\begin{aligned} \text{minimize}_{\boldsymbol{x}} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\,. \end{aligned}$$

Problem 8: Given $\boldsymbol{b} \in \mathbb{R}^M$, $\boldsymbol{A} \in \mathbb{R}^{M \times d}$ and positive definite $\boldsymbol{Q} \in \mathbb{R}^{d \times d}$, derive the Lagrange dual problem of

$$\begin{aligned} \text{minimize}_{\boldsymbol{x}} \quad & \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\,. \end{aligned}$$

4 Constrained Optimization: Toy Problem

Problem 9: Suppose we have 40 pieces of raw material. Toy A can be made of one piece material with 3 EUR machining fee. A larger toy B can be made from two pieces of material with 5 EUR machining fee.

Because distribution costs decrease with larger quantities, we can sell x pieces of toy A for $20 - x$ EUR each, and y pieces of toy B for $40 - y$ EUR each. From our experience, toy B is more popular than toy A; therefore, we will produce not more of toy A than of toy B.

To get the maximum profit, we want to calculate the amount of toy A and toy B that we should produce.

a) Write down the constrained optimization problem and the associated Lagrangian.

b) Write down the Karush–Kuhn–Tucker (KKT) conditions for the above optimization problem.

c) Obtain the solution to the constrained optimization problem by solving the KKT conditions. Do not worry about non-integer production quantities.