IEOR 169: Problem Set #2

Spring 2022

Due Feb 24 2022 at 2 PM PST

You are encouraged to collaborate with fellow students as you work through the problem set. However, your final submission must be your own work.

If you used any technology to solve a problem (Pyomo, Gurobipy, Excel, AMPL, etc..) make sure to include the relevant details (ipython notebook, pdf of excel set up and solution, pdf of AMPL code and output, etc..). I may request a copy of the original files that you used.

Optional problems will not push your grade beyond 100% but it may compensate for your mistakes in the main problem set.

Your solutions must be uploaded on bCourses by Feb 24 2022 at 2 PM PST. Late submissions without prior approval will not be accepted.

Problem 1 ML objective linearization

A classic problem of machine learning is the problem of a scalar regression when we are trying to learn a function $f_x(a) = b$ from a dataset $\{a_i, b_i\}_{i=1}^m$ where $a_i \in \mathbb{R}^d$ is a data entry and $b_i \in \mathbb{R}$ is the estimated value. Here $f : \mathbb{R}^d \to \mathbb{R}$ is a parametric function given by a vector of parameters $x \in \mathbb{R}^n$. Given the true value b and its estimation $\hat{b}$ the loss of $\ell(b, \hat{b}) \in \mathbb{R}$ is said to be incurred due to the error in estimation. Then, the classical problem of training the regression is defined as

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^m \ell(b_i, f_x(a_i))$$

For the purpose of this exercise, we assume the common choice of the loss function $\ell(b, \hat{b}) = |b - \hat{b}|$, also called the ℓ_1-loss or the absolute value loss (it is often used if there is sparse noise in the data, such as outliers). The problem of ℓ_1 regression can be formulated as

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^m |b_i - f_x(a_i)|$$

1. (Single-layer perceptron) Consider the function f_x given with the simplest version of an artificial neural network: the single-layer perceptron with a linear activation function (this is also called the linear model). Mathematically, it is be written as
$$f_x(a) = a^\top x$$
Write down the problem of ℓ_1 regression as a linear optimization problem

2. (Multy-layer perceptron) Consider the two-layer perceptron with the rectified linear unit (ReLU, $(x)_+ = \max\{0, x\}$) activation function. It is given in the form
$$f_{w^1,\dots,w^n,v}(a) = \sum_{j=1}^{l} v_j \max\{0, w^j a\}$$
Write down the problem of ℓ_1 regression as a mixed-integer linear optimization problem (note that the problem now consists in optimizing over vectors $v \in \mathbb{R}^l$ and $w^1, \dots, w^l \in \mathbb{R}^d$) You can assume that there is some number M that limits from above the absolute value of any other number in the problem.

. . . .

Problem 2 primal and dual bounds for integer knapsack

Exercise 2.9.3 from the textbook

. . . .

Problem 3 relaxations I

Exercise 2.9.4 from the textbook

. . . .

Problem 4 relaxations II

Exercise 2.9.5 from the textbook

. . . .

Problem 5 relaxations III

Exercise 2.9.8 from the textbook

. . . .

Problem 6 Lagrangian duality of knapsack

Exercise 10.7.5 from the textbook (part (i) only)

. . . .

Problem 7 surrogate duality (optional)

Exercise 10.7.9 from the textbook

. . . .