COMP 3105 – Assignment 1 Report
– Fall 2025 –

Due: Sunday September 28, 2025 23:59.

Group 51
Andrew Wallace - 101210291

Question 1: (7.5%) Linear Regression

(a) (1%) L_2 Regression
Please see A1codes.py for implementation.

(b) (3%) L_∞ Regression
Here we are going to solve the L_∞ loss regression problem

$$\mathbf{w} = \text{argmin}_{\mathbf{w}\in\mathbb{R}^d} = ||X\mathbf{w} - \mathbf{y}||$$

Recall that this optimization can be expressed as a linear programming problem with the joint paramters $\begin{bmatrix} \underset{d\times 1}{\mathbf{w}} \\ \underset{1\times 1}{\delta} \end{bmatrix} \in \mathbb{R}^{d+1}$ as follows

$$\begin{gathered} \min_{\mathbf{w},\delta}\delta \\ \text{s.t.} \\ \delta \geq 0 \iff -\delta \leq 0 \\ X\mathbf{w} - \mathbf{y} \preceq \delta \cdot \mathbf{1}_n \iff X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq \mathbf{y} \\ \mathbf{y} - X\mathbf{w} \preceq \delta \cdot \mathbf{1}_n \iff -X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq -\mathbf{y} \end{gathered}$$

In the following answers, we will convert the optimization to a form that is solvable by the **cxvopt** linear programming (LP) solver, which solves the following form of LP

$$\begin{gathered} \min_{\mathbf{u}}\mathbf{c}^T\mathbf{u} \\ \text{s.t. } G\mathbf{u} \preceq \mathbf{h} \end{gathered}$$

Let the unknown variables be $\mathbf{u} = \begin{bmatrix} \underset{d\times 1}{\mathbf{w}} \\ \underset{1\times 1}{\delta} \end{bmatrix} \in \mathbb{R}^{d+1}$

For the constraints, since we have three sets of constraints, the matrix G and $\mathbf{h}$ can be decomposed into three parts

$$G \cdot \mathbf{u} = \begin{bmatrix} \underset{1\times(d+1)}{G^{(1)}} \\ \underset{n\times(d+1)}{G^{(2)}} \\ \underset{n\times(d+1)}{G^{(3)}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{w} \\ \delta \end{bmatrix} \preceq \begin{bmatrix} \underset{1\times 1}{\mathbf{h}^{(1)}} \\ \underset{d\times 1}{\mathbf{h}^{(2)}} \\ \underset{d\times 1}{\mathbf{h}^{(3)}} \end{bmatrix}$$

(b.1) (0.25%) For the objective function, we want $\mathbf{c}^T\mathbf{u} = \delta$. What should $\mathbf{c} \in \mathbb{R}^{d+1}$ be?
Recall that $\mathbf{u} = \begin{bmatrix} \mathbf{w} \\ \delta \end{bmatrix}$
For the objective function to be δ we have:

$$\mathbf{c}^T\mathbf{u} = [c_1, c_2, \dots c_d, c_{d+1}] \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \\ \delta \end{bmatrix} = c_1w_1 + c_2w_2 + \ldots + w_du_d + c_{d+1}\delta$$

Now let $c_1 = c_2 = \ldots = c_d = 0$
And $c_{d+1} = 1$
This gives us:
$0w_1 + 0w_2 + \ldots 0w_d + \delta$
$= \delta.$

Thus $\mathbf{c} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$
Where $c_1 = c_2 = \ldots = c_d = 0$ and $c_{d+1} = 1$.

(b.2) (0.25%) We want $G^{(1)}\mathbf{u} \preceq \mathbf{h}^{(1)} \iff \delta \geq 0$. What should $G^{(1)} \in \mathbb{R}^{1\times(d+1)}$ and $h^{(1)} \in \mathbb{R}$ be?
Recall the first constraint: $-\delta \leq 0$
We want

$$\underset{1\times(d+1)}{G^{(1)}} \cdot \mathbf{u} \preceq \underset{1\times 1}{\mathbf{h}^{(1)}}$$

We will now map our constraint into this form.

$$\begin{bmatrix} \underset{1\times d}{G^{(11)}} & \underset{1\times 1}{G^{(12)}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{w} \\ \delta \end{bmatrix} \preceq \underset{n\times 1}{\mathbf{h}^{(1)}} \iff -\delta \leq 0$$
$$\underset{1\times d}{G^{(11)}} \cdot \mathbf{w} + \underset{1\times 1}{G^{(12)}} \cdot \delta = -\delta \leq 0 = \underset{1\times 1}{\mathbf{h}^{(1)}}$$

So,
$\underset{1\times d}{G^{(11)}} = \mathbf{0}_{1\times d}$
$\underset{1\times 1}{G^{(12)}} = -1$
$\underset{1\times(d+1)}{G^{(1)}} = \begin{bmatrix} \mathbf{0}_{1\times d} & -1 \end{bmatrix}$
$\underset{1\times 1}{\mathbf{h}^{(1)}} = \mathbf{0}_{1\times 1}$

(b.3) (0.25%) We want $G^{(2)}\mathbf{u} \preceq \mathbf{h}^{(2)} \iff X\mathbf{w} - \mathbf{y} \preceq \delta \cdot \mathbf{1_n}$. What should $G^{(2)} \in \mathbb{R}^{n\times(d+1)}$ and $h^{(2)} \in \mathbb{R}^n$ be?
Recall our second constraint is
$X\mathbf{w} - \mathbf{y} \preceq \delta \cdot \mathbf{1}_n \iff X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq \mathbf{y}$
We want

$$\underset{n\times(d+1)}{G^{(2)}} \cdot \mathbf{u} \preceq \underset{n\times 1}{\mathbf{h}^{(2)}}$$

We will now map our constraint into this form.

$$\begin{bmatrix} \underset{n\times d}{G^{(21)}} & \underset{n\times 1}{G^{(22)}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{w} \\ \delta \end{bmatrix} \preceq \underset{n\times 1}{\mathbf{h}^{(2)}} \iff X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq \mathbf{y}$$
$$\underset{n\times d}{G^{(21)}} \cdot \mathbf{w} + \underset{n\times 1}{G^{(22)}} \cdot \delta = X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq \mathbf{y} = \underset{n\times 1}{\mathbf{h}^{(1)}}$$

So,
$\underset{n\times d}{G^{(21)}} = X$
$\underset{n\times 1}{G^{(22)}} = -\mathbf{1}_{n\times 1}$
$\underset{n\times (d+1)}{G^{(2)}} = \begin{bmatrix} X & -\mathbf{1}_{n\times 1} \end{bmatrix}$
$\underset{n\times 1}{h^{(2)}} = \mathbf{y}$

(b.4) (0.25%) We want $G^{(3)}\mathbf{u} \preceq \mathbf{h}^{(3)} \iff \mathbf{y} - X\mathbf{w} \preceq \delta \cdot \mathbf{1_n}$. What should $G^{(3)} \in \mathbb{R}^{n\times(d+1)}$ and $h^{(3)} \in \mathbb{R}^n$ be?
Recall our third constraint is
$\mathbf{y} - X\mathbf{w} \preceq \delta \cdot \mathbf{1}_n \iff -X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq -\mathbf{y}$
We want

$$\underset{n\times (d+1)}{G^{(3)}} \cdot \mathbf{u} \preceq \underset{n\times 1}{\mathbf{h}^{(3)}}$$

We will now map our constraint into this form.

$$\begin{bmatrix} \underset{n\times d}{G^{(31)}} & \underset{n\times 1}{G^{(32)}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{w} \\ \delta \end{bmatrix} \preceq \underset{n\times 1}{\mathbf{h}^{(3)}} \iff -X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq -\mathbf{y}$$
$$\underset{n\times d}{G^{(31)}} \cdot \mathbf{w} + \underset{n\times 1}{G^{(32)}} \cdot \delta = -X\mathbf{w} - \delta \cdot \mathbf{1}_n \preceq -\mathbf{y} = \underset{n\times 1}{\mathbf{h}^{(3)}}$$

So,
$\underset{n\times d}{G^{(31)}} = -X$
$\underset{n\times 1}{G^{(32)}} = -\mathbf{1}_{n\times 1}$
$\underset{n\times (d+1)}{G^{(3)}} = \begin{bmatrix} -X & -\mathbf{1}_{n\times 1} \end{bmatrix}$
$\underset{n\times 1}{h^{(3)}} = -\mathbf{y}$

(b.5) (2%) Based on your derivations in (b), implement a Python function

$$\mathbf{w} = \text{minimizeLinf}(X, y)$$

that returns a $d \times 1$ vector of weights/parameters $\mathbf{w}$ corresponding to the solution of minimum L_∞ loss.

$$\mathbf{w} = \text{argmin}_{w\in\mathbb{R}^d} ||X\mathbf{w} - \mathbf{y}||$$

Based on our derivations in (b), we result in

$$\mathbf{c} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^{d+1}$$

$$\mathbf{u} = \begin{bmatrix} \mathbf{w} \\ \delta \end{bmatrix} \in \mathbb{R}^{d+1}$$

$$G = \begin{bmatrix} \mathbf{0}_{1\times d} & -1 \\ X & -\mathbf{1}_n \\ -X & -\mathbf{1}_n \end{bmatrix} \in \mathbb{R}^{(2n+1)\times(d+1)}$$

$$\mathbf{h} = \begin{bmatrix} 0 \\ \mathbf{y} \\ -\mathbf{y} \end{bmatrix} \in \mathbb{R}^{2n+1}$$

$$\text{s.t.} \quad \underset{(2n+1)\times(d+1)}{G} \cdot \underset{(d+1)\times 1}{\mathbf{u}} \preceq \underset{(2n+1)\times 1}{\mathbf{h}}$$

Please see A1codes.py for full implementation.

(c) (2%) Synthetic Regression Problem
In this part, you will evaluate your implemented algorithms on a synthetic dataset.

(c.1) (1%) Implement a Python function
`train_loss, test_loss = synRegExperiments()`
that returns a 2 x 2 matrix train loss of average training losses and a 2 x 2 matrix test loss of average test losses (See Table 1 and Table 2 below.) It repeats 100 runs as follows
Please see **A1codes.py** for full details.

(c.2) (1%) Looking at your tables from above, analyze the results and discuss any findings you may have and the possible reason behind them.
Note that the matrices below are representative of the following table:

Model	Average L_2 loss	Average L_∞ loss
L_2 model L_∞ model		

Table 1: Different training losses for different models

After running `synRegExperiments` on the regression_train.csv and regression_test.csv data sets we get the following results:
For training data:

$$\begin{bmatrix} 2.00305568 & 7.31711154 \\ 2.36347873 & 6.6215357 \end{bmatrix}$$

For test data:
$$\begin{bmatrix} 1.66071948 & 5.58874693 \\ 1.9868478 & 5.02664782 \end{bmatrix}$$

For a randomly generated training and test data set using seed 101210291, we get the following results:

For training data:
$$\begin{bmatrix} 0.11160734 & 1.68718399 \\ 0.27832791 & 0.91119205 \end{bmatrix}$$
For test data:
$$\begin{bmatrix} 0.05252988 & 1.02181983 \\ 0.35054548 & 2.26890905 \end{bmatrix}$$

This shows that our L_2 loss, overall, has a lower average loss compared to the L_∞ loss.