Name: Hongda Li Class: CSE 546 HW2B

B.1

Objective: Given the definition for the L2, L1 and the Infinity norm of real vector, show that $\|x\|_\infty \le \|x\|_2 \le \|x\|_1$.
First we are going to show that $\|x\|_2^2 \le \|x\|_1^2$, starting from the definition of the norms we have:

$$\begin{aligned}\|x\|_1^2 &= \left(\sum_{i=1}^{n}|x_i|\right)^2 \qquad \text{(B.1.1)}\\ &= \sum_{i=1}^{n}\left(|x_i|\sum_{j=1}^{n}|x_j|\right)\\ &= \sum_{i=1}^{n}\left(|x_i|^2+|x_i|\sum_{j=1,j\neq i}^{n}|x_j|\right)\\ &= \sum_{i=1}^{n}|x_i|^2+\sum_{i=1}^{n}|x_i|\left(\sum_{j=1,j\neq i}^{n}|x_j|\right)\\ &= \|x\|_2^2+\underbrace{\sum_{i=2}^{n}\sum_{j=1}^{i-1}2|x_i||x_j|}_{\ge 0}\\ \implies &\|x\|_2^2 \le \|x\|_1^2\end{aligned}$$

And now we are going to shoe that $\|x\|_\infty^2 \le \|x\|_2^2$. By the definition of the infinity norm, we know that therde exists $1 \le m \le n$ such that $x_m = \|x\|_\infty = \max_{1\le i\le n}(x_i)$. Then it can be said that:

$$\begin{aligned}x_m^2 \le& x_m^2 + \underbrace{\sum_{i=1,i\neq m}^{n} x_i^2}_{\ge 0} \qquad \text{(B.1.2)}\\ x_m^2 =& \|x\|_\infty \le \sum_{i=1}^{n}x_i^2 = \|x\|_2^2\end{aligned}$$

And then combing together, we can take the square root because the function $\sqrt{\bullet}$ is monotone increase, hence it preserves the inequality, which will give us $\|x\|_\infty \le \|x\|_2^2 \le \|x\|_1$.

B.2

B.2.a

Objective: The function $\|x\|$ is a convex function.

$$\begin{aligned}\|\lambda x+(1-\lambda)y\| &\le \|\lambda x\| + \|(1-\lambda)y\| \qquad \text{(B.2.a.1)}\\ &= \lambda\|x\| + (1-\lambda)\|y\|\end{aligned}$$

Note, I just directly apply the Triangular inequality of the norm to get the inequality, and then becaues $\lambda \in [0,1]$, so there is no absolute value, and notice that the resulting expression is the definition of Convexity the given function.

B.2.b

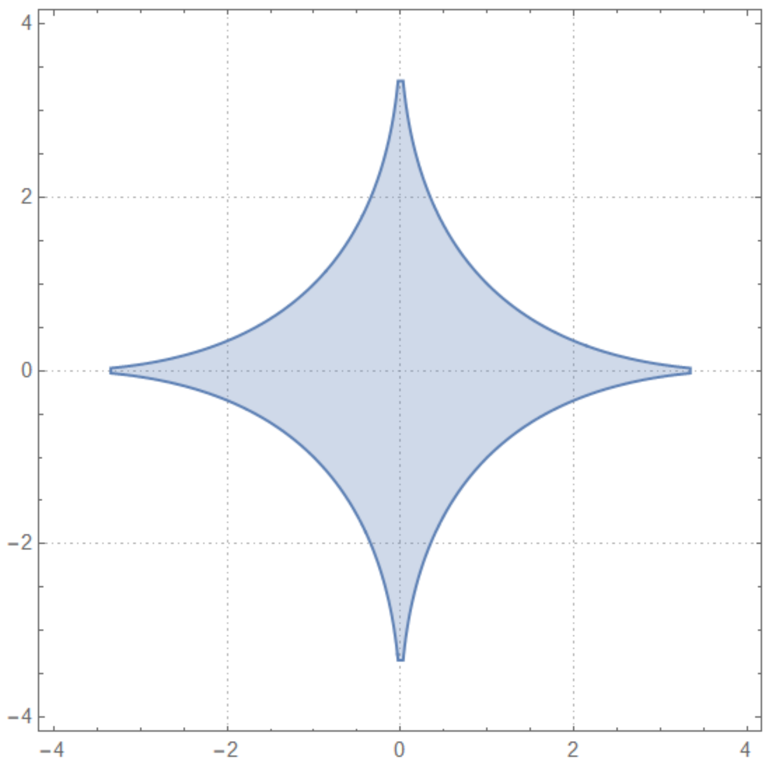
Objective: Show that the set $\{x \in \mathbb{R}^n : \|x\| \leq 1\}$ is a convex set. Let the set be denoted as S Let's take any 2 points in the set like $x \in S$, $y \in s$, then $\|x\| \leq 1$ and $\|y\| \leq 1$ for any line defined by the 2 points:

$$\|\lambda x + (1-\lambda)y\| \leq \lambda \underbrace{\|x\|}_{\leq \lambda} + \underbrace{(1-\lambda)\|y\|}_{\leq 1-\lambda} \tag{B.2.b.1}$$
$$\implies \|\lambda x + (1-\lambda)y\| \leq 1$$
$$\implies \lambda x + (1-\lambda)y \in S$$

The first by the inequality of norm, and the second is by the definition of the fact that $x, y \in S$, and the third is by the definition of the set S.

B.2.c

The set $\{(x_1, x_2) : g(x_1, x_2) \leq 4\}$ domain of the function such that the value of the function is bounded by a given quantity. This is the unit norm ball defined by $p = \frac{1}{2}$. As we showned in lecture, it's not convex, and it looks like a star[1]:



B.3

B.3.a

Objective: Showing that, the squared loss function regularized with Eclidean norm (not squared) is convex. The trick is to show that, the squared loss function is convex, and we have shown the the Euclidean norm is convex back in B.2.a. And by showing that the positive weighted sum of 2 convex function is convex, we will be able to show that it's convex.

The loss function is convex because: $l_i(w) = (y_i - w^T x_i)^2$, this function is qaudratic, and it's second derivative

[1]Drawn by mathematica.

is diag(2) (The hessian fof the function wrt to w), which is a positive quantity. When the second derivative of a function is always larger than or equals to zero, the function is convex.
Here, we will show that the sum of 2 convex function is convex too. Let $f(x), g(x)$ be 2 convex function mapping $x \in \mathbb{R}^n$ to $\mathbb{R}$. Choose any 2 points $x, y \in \mathbb{R}^n$ then we have:

$$
\begin{aligned}
& f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \qquad \text{(B.3.a.1)} \\
& \wedge \\
& g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y) \\
\implies & f(\lambda x + (1-\lambda)y) + g(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) + \lambda g(x) + (1-\lambda)g(y) \\
& f(\lambda x + (1-\lambda)y) + g(\lambda x + (1-\lambda)y) \leq \lambda(f(x) + g(x)) + (1-\lambda)(f(x) + g(x))
\end{aligned}
$$

It's not hatd to see that, if we set $h(x) := g(x) + h(x)$ then the last line of the statement will be the convexity of $h(x)$, proving the that sume of the 2 functions are still convex.
Note: If the sum 2 functions are still convex, then **the sum finite many function is still going to be convex**, this can be proved inductively, because the + sign is associatitve.
Using this, and using the fact that $\sum_{i=1}^{n} l_i(w) + \lambda\|w\|$ is convex, because all the function we are summing up are convex function.

B.2.b

Convex loss function is easier to optimize, it has minimal, and it's a convex set of minimal as well. And the use of a **strictly convex regularizer** will give **unique global optimal**, the strongly convex loss function is the best, because gradient descent performs really well on them.

B.4: Multinomial Logistic Regression

B.5: Confidence Interval of Least Squares Estimation: Bounding Estimate