University of Waterloo
CS 480
r2knowle: 2023-11-28

Exercise # 2

Q2) To begin we are given the following inequality to try and minimize:

$$\min_{w\in\mathbb{R}} \max_{\forall j, ||z_j|| \leq \lambda} ||(X+Z)w - y||_2$$

We can expand this to get:

$$\min_{w\in\mathbb{R}} \max_{\forall j, ||z_j|| \leq \lambda} ||Xw - y + Zw||_2$$

Going forward we will try to prove that this results in (3) by using the squeeze theorem. We will first prove that the upper bound results in (3). By the triangle identity it follows that: Now by triangle inequality it follows that:

$$\min_{w\in\mathbb{R}} \max_{\forall j, ||z_j|| \leq \lambda} ||Xw - y + Zw||_2 \leq \min_{w\in\mathbb{R}} \max_{\forall j, ||z_j|| \leq \lambda} (||Xw - y||_2 + ||Zw||_2)$$

This can further be reduced to:

$$\leq \min_{w\in\mathbb{R}} \left(||Xw - y||_2 + \max_{\forall j, ||z_j|| \leq \lambda} || \sum_j^d (z_j w_j) ||_2 \right)$$

from our inner maximization we can further get that:

$$\leq \min_{w\in\mathbb{R}} \left(||Xw - y||_2 + || \sum_j^d (\lambda w_j) ||_2 \right)$$

$$= \min_{w\in\mathbb{R}} \left(||Xw - y||_2 + ||\lambda \sum_j^d (w_j) ||_2 \right)$$

From the homogeneous nature of the l_2 norm we know that this is equivalent to:

$$\leq \min_{w\in\mathbb{R}} \left(||Xw - y||_2 + |\lambda| || \sum_j^d (w_j) ||_2 \right)$$

Since λ is a positive value, it follows that $|\lambda| = \lambda$, moreover for any $w \in R^d$ it follows that $||w||_2 \leq ||w||_1$. This is because $||w||_2 = \max_{||u||_2 \leq 1} w^T u$ and $||w||_1 = |w| > w^T u$ for any u thats less then 1. Therefore it follows that:

$$\leq \min_{w\in\mathbb{R}} \left(||Xw - y||_2 + \lambda || \sum_j^d (w_j) ||_1 \right)$$

$$= \min_{w\in\mathbb{R}} (||Xw - y||_2 + \lambda ||w||_1)$$

Thus showing how we can get (3) from (2) as required in the upper bound case. On the next page we are going to be looking at the lower bound case.

To prove the lower bound, we will start with:

$$\min_{w \in \mathbb{R}} \max_{\forall j, ||z_j|| \leq \lambda} ||Xw - y + Zw||_2$$

We will define the following variable that has the properties:

$$h = \begin{cases} \frac{Xw-y}{||Xw-y||_2} & Xw \neq y \\ 1 & Xw = y \end{cases}$$

We then make the observation that:

$$\forall j, ||\lambda||_2 \geq ||\lambda h||_2$$

Therefore if we will set z_j to be the following:

$$\tilde{z_j} = \lambda h$$

This has the property that it will always be less then or equal to our maximal quantity, thus we can rewrite this as:

$$\min_{w \in \mathbb{R}} \max_{\forall j, ||z_j|| \leq \lambda} ||(Xw - y + Zw||_2 \geq \min_{w \in \mathbb{R}} ||Xw - y + \sum_j^d \tilde{z}_j w_j||_2$$

Which from this we can get:

$$= \min_{w \in \mathbb{R}} ||Xw - y + \sum_j^d \lambda h w_j||_2$$

$$= \min_{w \in \mathbb{R}} ||Xw - y + h \sum_j^d \lambda |w_j|||_2$$

$$= \min_{w \in \mathbb{R}} ||Xw - y||_2 + \sum_j^d \lambda |w_j|$$

$$= \min_{w \in \mathbb{R}} ||Xw - y||_2 + \lambda |w|$$

Thus demonstrating how we can achieve (3) in the lower bound case. Thus because we have shown the lower bound and upper bound of (2) converge to (3) it must be the case that (2) is equivalent to (3).