

# CS754 Assignment 3 Report

Arpon Basu  
Shashwat Garg

Spring 2022

## Contents

<b>1</b>	<b>Problem 1</b>	<b>2</b>
<b>2</b>	<b>Problem 2</b>	<b>2</b>
<b>3</b>	<b>Problem 3</b>	<b>4</b>
<b>4</b>	<b>Problem 4</b>	<b>4</b>
4.1	Introduction and Aim . . . . .	4
4.2	Mathematical Formulation . . . . .	4
<b>5</b>	<b>Problem 5</b>	<b>4</b>
<b>6</b>	<b>Problem 6</b>	<b>5</b>

## Introduction

Welcome to our report on CS754 Assignment 3. We have tried to make this report comprehensive and self-contained. We hope reading this would give you a proper flowing description of our work, methods used and the results obtained.

Hope you enjoy reading the report. Here we go!

## 1 Problem 1

## 2 Problem 2

(a) The restricted eigenvalue property has a very neat and succinct description in terms of the SVD of our sensing matrix  $\mathbf{X}$ , and that's how we'll define it.

Note that the restricted eigenvalue property for the matrix  $X$  is defined as follows

$$\frac{1}{N \cdot \|\nu\|_2^2} \nu^T X^T X \nu \geq \gamma \quad \forall \nu \in \mathbb{K}^{p \times 1} \setminus \{\mathbf{0}\}$$

where  $\gamma$  is the coefficient of the property.

Now, notice that if we restrict  $\nu$  to vary over only unit length vectors in  $\mathbb{K}^{p \times 1}$ , then we can get rid of the  $\|\nu\|_2^2$  in the denominator. Also, since the inequality holds for all  $\nu$  in our chosen space, it's enough to assume that the infimum of the expression  $\frac{1}{N} \nu^T X^T X \nu$  over all  $\nu$  such that  $\|\nu\|_2 = 1$  must be greater than or equal to  $\gamma$ , ie:-

$$\begin{aligned} \min_{\|\nu\|_2=1} \frac{1}{N} \nu^T X^T X \nu &\geq \gamma \\ \implies \min_{\|\nu\|_2=1} (X\nu)^T X \nu &\geq N\gamma \\ \implies \min_{\|\nu\|_2=1} \|X\nu\|_2^2 &\geq N\gamma \\ \implies \min_{\|\nu\|_2=1} \|X\nu\|_2 &\geq \sqrt{N\gamma} \end{aligned}$$

But from the theory of Singular Value Decomposition, we know that  $\min_{\|\nu\|_2=1} \|X\nu\|_2$  is equal to the smallest singular value of  $X$ .

Thus, we can **define** the **Restricted Eigenvalue Property** with coefficient  $\gamma$  as follows : The matrix  $\mathbf{X} \in \mathbb{R}^{n \times p}$  is said to satisfy the Restricted Eigenvalue Property with coefficient  $\gamma$  if the **smallest singular value of  $X$  is atleast  $\sqrt{N\gamma}$** .

(b) We have that

$$\begin{aligned} G(\nu) &:= \frac{1}{2N} \|y - X(\beta^* + \nu)\|_2^2 + \lambda_N \|\beta^* + \nu\|_1 \\ J(\mathbf{x}) &:= \frac{1}{2N} \|y - X\mathbf{x}\|_2^2 + \lambda_N \|\mathbf{x}\|_1 \end{aligned}$$

Thus

$$G(0) := \frac{1}{2N} \|y - X\beta^*\|_2^2 + \lambda_N \|\beta^*\|_1 = J(\beta^*)$$

On the other hand, for  $\hat{\nu} := \hat{\beta} - \beta^*$ , we have

$$G(\hat{\nu}) := \frac{1}{2N} \|y - X\hat{\beta}\|_2^2 + \lambda_N \|\hat{\beta}\|_1 = J(\hat{\beta})$$

But  $\hat{\beta}$  was, **by construction** (as mentioned on Pg# 308, third last line), the optimizer of the LASSO function  $J(\mathbf{x})$ , and consequently  $G(\hat{\nu}) = J(\hat{\beta}) \leq J(\beta^*) = G(0)$ , as desired. (c) We just have to piece together some equations and inequalities to derive equation 11.21. They go as follows:

$$G(\hat{\nu}) \leq G(0)$$

$$\mathbf{y} = \mathbf{X}\beta^* + \mathbf{w}$$

$$\hat{\beta} = \beta^* + \hat{\nu}$$

The first inequality  $G(\hat{\nu}) \leq G(0)$ , combined with the fact that  $\hat{\beta} = \beta^* + \hat{\nu}$  yields

$$G(\hat{\nu}) \leq G(0)$$

$$\implies \frac{1}{2N} \|\mathbf{y} - \mathbf{X}\hat{\beta}\|_2^2 + \lambda_N \|\hat{\beta}\|_1 \leq \frac{1}{2N} \|\mathbf{y} - \mathbf{X}\beta^*\|_2^2 + \lambda_N \|\beta^*\|_1$$

$$\implies \frac{1}{2N} \|\mathbf{y} - \mathbf{X}(\beta^* + \hat{\nu})\|_2^2 + \lambda_N \|\beta^* + \hat{\nu}\|_1 \leq \frac{1}{2N} \|\mathbf{y} - \mathbf{X}\beta^*\|_2^2 + \lambda_N \|\beta^*\|_1$$

Now, using the fact that  $\mathbf{y} = \mathbf{X}\beta^* + \mathbf{w}$ , we get

$$\frac{1}{2N} \|\mathbf{y} - \mathbf{X}(\beta^* + \hat{\nu})\|_2^2 + \lambda_N \|\beta^* + \hat{\nu}\|_1 \leq \frac{1}{2N} \|\mathbf{y} - \mathbf{X}\beta^*\|_2^2 + \lambda_N \|\beta^*\|_1$$

$$\implies \frac{1}{2N} \|\mathbf{y} - \mathbf{X}\beta^* - \mathbf{X}\hat{\nu}\|_2^2 + \lambda_N \|\beta^* + \hat{\nu}\|_1 \leq \frac{1}{2N} \|\mathbf{w}\|_2^2 + \lambda_N \|\beta^*\|_1$$

$$\implies \frac{1}{2N} \|\mathbf{w} - \mathbf{X}\hat{\nu}\|_2^2 \leq \frac{1}{2N} \|\mathbf{w}\|_2^2 + \lambda_N (\|\beta^*\|_1 - \|\beta^* + \hat{\nu}\|_1)$$

Now,  $\|\mathbf{w} - \mathbf{X}\hat{\nu}\|_2^2 = (\mathbf{w} - \mathbf{X}\hat{\nu})^T \cdot (\mathbf{w} - \mathbf{X}\hat{\nu}) = \|\mathbf{w}\|_2^2 + \|\mathbf{X}\hat{\nu}\|_2^2 - 2\mathbf{w}^T \mathbf{X}\hat{\nu}$ . Substituting this in the equation above yields

$$\frac{1}{2N} (\|\mathbf{w}\|_2^2 + \|\mathbf{X}\hat{\nu}\|_2^2 - 2\mathbf{w}^T \mathbf{X}\hat{\nu}) \leq \frac{1}{2N} \|\mathbf{w}\|_2^2 + \lambda_N \{\|\beta^*\|_1 - \|\beta^* + \hat{\nu}\|_1\}$$

$$\implies \frac{1}{2N} (\|\mathbf{X}\hat{\nu}\|_2^2 - 2\mathbf{w}^T \mathbf{X}\hat{\nu}) \leq \lambda_N \{\|\beta^*\|_1 - \|\beta^* + \hat{\nu}\|_1\}$$

$$\implies \frac{1}{2N} \|\mathbf{X}\hat{\nu}\|_2^2 - \frac{1}{N} \mathbf{w}^T \mathbf{X}\hat{\nu} \leq \lambda_N \{\|\beta^*\|_1 - \|\beta^* + \hat{\nu}\|_1\}$$

$$\implies \frac{\|\mathbf{X}\hat{\nu}\|_2^2}{2N} \leq \frac{\mathbf{w}^T \mathbf{X}\hat{\nu}}{N} + \lambda_N \{\|\beta^*\|_1 - \|\beta^* + \hat{\nu}\|_1\}$$

as desired.

(d) We use a version of Hölder's inequality for vectors as follows:

Let  $p, q \in [1, \infty]$  be two real numbers such that  $\frac{1}{p} + \frac{1}{q} = 1$ . Let  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  be two vectors. Then

$$\sum_{k=1}^n |\mathbf{x}_k \mathbf{y}_k| \leq \left( \sum_{k=1}^n |\mathbf{x}_k|^p \right)^{\frac{1}{p}} \left( \sum_{k=1}^n |\mathbf{y}_k|^q \right)^{\frac{1}{q}} = \|\mathbf{x}\|_p \|\mathbf{y}\|_q$$

We augment this inequality for our purposes: Note that  $\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^n x_k y_k \leq \sum_{k=1}^n |x_k y_k| \leq \left( \sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left( \sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}}$ , thus showing that  $\mathbf{x} \cdot \mathbf{y} \leq \left( \sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left( \sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}} = \|\mathbf{x}\|_p \|\mathbf{y}\|_q$ .

In our context let  $\mathbf{x} = X^T w$  and  $\mathbf{y} = \nu$ , and let  $p = \infty$ ,  $q = 1$ . Then,  $(X^T w) \cdot \nu = (X^T w)^T \nu = w^T X \nu$ . Then we have that

$$\begin{aligned} (X^T w) \cdot \nu &\leq \|X^T w\|_\infty \|\nu\|_1 \\ \implies w^T X \nu &\leq \|X^T w\|_\infty \|\nu\|_1 \end{aligned}$$

Thus,

$$\frac{w^T X \hat{\nu}}{N} + \lambda_N \{\|\hat{\nu}_S\|_1 - \|\hat{\nu}_{S^c}\|_1\} \leq \frac{\|X^T w\|_\infty}{N} \|\nu\|_1 + \lambda_N \{\|\hat{\nu}_S\|_1 - \|\hat{\nu}_{S^c}\|_1\}$$

as desired.

### 3 Problem 3

### 4 Problem 4

Paper Details	
Title of the Paper	Coastal Acoustic Tomography System and Its Field Application
Link of the paper	<a href="#">Click Here</a>
Author List	Haruhiko Yamoaka, Arata Kaneko, Jae-Hun Park, Hong Zheng, Noriaki Gohda, Tadashi Takano, Xiao-Hua Zhu and Yoshio Takasugi
Publication Date	August 2002
Publication Venue	IEEE Journal of Oceanic Engineering, Volume 27, Issue 2

#### 4.1 Introduction and Aim

This paper aims to map the structure of the “strongly nonlinear tidal currents in the coastal sea” by using multiple synchronised coastal acoustic tomography system (CATS). Using GPS clock signals and separate codes to distinguish between signals of individual systems, reconstruction of tidal process behaviour is done through an inverse analysis of the acoustic signals obtained by the sensors.

#### 4.2 Mathematical Formulation

### 5 Problem 5

We know that Radon Transform is given by-

$$R_\theta(f) = g(\rho, \theta) = \int_{-\infty}^{+\infty} f(\rho \cos \theta - z \sin \theta, \rho \sin \theta + z \cos \theta) dz$$

We can write the same as-

$$R_\theta(f) = g(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy$$

Let the scaled image be denoted by  $h(x, y) = f(ax, ay)$ . This is the same image as original, but scaled by a factor of  $a$ , in both  $x$  and  $y$  directions.

We can write the same Radon Transform as-

$$\begin{aligned} R_\theta(h) &= g'(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy \\ R_\theta(h) &= g'(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(ax, ay) \delta(x \cos \theta + y \sin \theta - \rho) dx dy \\ R_\theta(h) &= g'(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x', y') \delta\left(\frac{x' \cos \theta + y' \sin \theta - a\rho}{a}\right) \frac{dx'}{a} \frac{dy'}{a} \end{aligned}$$

Since  $\delta(ax) = \delta(x)/a$ , we get-

$$\begin{aligned} R_\theta(h) &= g'(\rho, \theta) = \frac{1}{a} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x', y') \delta(x' \cos \theta + y' \sin \theta - a\rho) dx' dy' \\ R_\theta(h) &= g'(\rho, \theta) = \frac{1}{a} g(a\rho, \theta) \end{aligned}$$

Thus, we can see that the Radon transform of the scaled image is also scaled by a factor of  $a$  in the size of projection, but the intensity of each projection has reduced by  $a$  as well.

## 6 Problem 6

We know that the Radon Transform is given by-

$$R_\theta(f) = g(\rho, \theta) = \int_{-\infty}^{+\infty} f(\rho \cos \theta - z \sin \theta, \rho \sin \theta + z \cos \theta) dz$$

We can write the same as-

$$R_\theta(f) = g(\rho, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy$$

Now, let  $f(x, y) = \delta(x - x_0, y - y_0)$  for some given constants  $x_0, y_0$ . Also, for the sake of simplification, call  $\delta(x \cos \theta + y \sin \theta - \rho)$  as  $h(x, y, \rho, \theta)$ .

Then the Radon transform of our function  $f(x, y)$  becomes

$$\begin{aligned} &\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x - x_0, y - y_0) h(x, y, \rho, \theta) dx dy \end{aligned}$$

Now, by a well known property of delta functions, the integration of a function multiplied by the delta function over any space (including the delta function's singularity) yields the

evaluation of the function at the singularity point. In the context of our problem, we can state the above property as

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x - x_0, y - y_0) h(x, y) dx dy = h(x_0, y_0)$$

Applying this property verbatim on our integral above yields

$$\begin{aligned} & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x - x_0, y - y_0) h(x, y, \rho, \theta) dx dy \\ &= h(x_0, y_0, \rho, \theta) = \delta(x_0 \cos \theta + y_0 \sin \theta - \rho) \end{aligned}$$

since  $\rho, \theta$  are constant parameters within the integration.

Thus the Radon transform of the unit impulse function is another impulse function, ie:-

$$R_{\theta}(\delta(x - x_0, y - y_0)) = g(\rho, \theta) = \delta(x_0 \cos \theta + y_0 \sin \theta - \rho)$$

$$R_{\theta}(\delta(x, y)) = g(\rho, \theta) = \delta(-\rho) = \delta(\rho)$$