CS 754 Assignment 3 : Report

Mls Pragna - 180050064
Rahul Puli - 180050080

Question 1

To run the code : ”run Q1.m”
RMSE value = 0.4157

original image

noise image

reconstructed image

Question 2

a)

The restricted eigenvalue condition is lower bounding the restricted eigenvalues of the model matrix,

$$\frac{\frac{1}{N} v^T X^T X v}{||v||_2^2} \geq \gamma \ \ \forall \text{ non-zero } v \in C$$

As our loss function $f_N(\beta) = \frac{1}{2N}||y - X\beta||_2^2$, $\beta \epsilon R^p$ is not strongly convex, we relax the strong convexity notion and make it necessary for some subset $C \epsilon R^p$. In particular, we say that a function function f satisfies strong convexity at β^* w.r.t to C if there is a constant $\gamma > 0$ such that

$$\frac{v^T \nabla^2 f(\beta) v}{||v||_2^2} \geq \gamma \ \ \forall \text{ non-zero } v \in C$$

In the specific case of linear regression, this notion is equivalent to lower bounding the restricted eigenvalues of the model matrix.

b)

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

As, $\hat{\nu} = \hat{\beta} - \beta^*$

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

Putting v=0 we get,

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

As, β^* is the true coefficient signal, $\hat{\beta}$ is the estimated signal and the minimum value of G(v) is achieved when $\nu = \hat{\nu}$

$$||\boldsymbol{y} - \boldsymbol{X}\hat{\beta}||_2^2 \leq ||\boldsymbol{y} - \boldsymbol{X}\beta^*||_2^2 \;\; and \;\; ||\hat{\beta}||_1 \leq ||\beta^*||_1$$

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

c)

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

As, $y = X\beta^* + w$,

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

$$\begin{aligned} G(0) &= \frac{1}{2N}||y - X\beta^*||_2^2 + \lambda_N||\beta^*||_1 \\ &= \frac{1}{2N}||X\beta^* + w - X\beta^*||_2^2 + \lambda_N||\beta^*||_1 \\ &= \frac{1}{2N}||w||_2^2 + \lambda_N||\beta^*||_1 \end{aligned}$$

Using the inequality $G(\hat{\nu}) \leq G(0)$ from part (b),

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

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

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

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

d)

To show, $\frac{||X\hat{\nu}||_2^2}{2N} \leq \frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1)$

By holder's inequality, the vectors u,v satisfy the below inequality

$$u^Tv \leq ||u||_p||v||_q$$

where p,q ϵ [1, ∞] with $\frac{1}{p} + \frac{1}{q} = 1$
Putting $u = X^Tw$, $v = \hat{\nu}$, $p = \infty$ and $q = 1$ in holder's inequality we get,

$$w^TX\hat{\nu} \leq ||X^Tw||_\infty||\hat{\nu}||_1 => \textbf{equation (1)}$$

As, β^* is a S sparse vector, $\beta^*_{S^c} = 0$,

$$||\beta^* + \hat{\nu}||_1 = ||\beta^*_S + \hat{\nu}_S||_1 + ||\beta^*_{S^c} + \hat{\nu}_{S^c}||_1 = ||\beta^*_S + \hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1$$

By triangle's inequality, $||a+b||_1 \geq |||a||_1 - ||b||_1|$,

$$||\beta^*_S||_1 - ||\hat{\nu}_S||_1 \leq ||\beta^*_S + \hat{\nu}_S||_1$$

$$\Rightarrow ||\beta^*_S||_1 - ||\hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1 \leq ||\beta^*_S + \hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1$$
$$\Rightarrow ||\beta^*_S||_1 - ||\hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1 \leq ||\beta^* + \hat{\nu}||_1$$
$$\Rightarrow ||\beta^*_S||_1 - ||\beta^* + \hat{\nu}||_1 \leq ||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1$$
$$\Rightarrow ||\beta^*||_1 - ||\beta^* + \hat{\nu}||_1 \leq ||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1 => \textbf{equation (2)}$$

Combining equation(1) and equation (2),

$$\frac{1}{N}w^TX\hat{\nu} + \lambda_N(||\beta^*||_1 - ||\beta^* + \hat{\nu}||_1) \leq \frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) => \textbf{equation (3)}$$

Combining inequality from part (c) $\frac{1}{2N}||X\hat{\nu}||_2^2 \leq \frac{1}{N}w^TX\hat{\nu} + \lambda_N(||\beta^*||_1 - ||\beta^* + \hat{\nu}||_1)$ and equation (3),

$$\frac{||X\hat{\nu}||_2^2}{2N} \leq \frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1)$$

e)

To show, $\frac{||\boldsymbol{X}\hat{\nu}||_2^2}{2N} \leq \frac{\lambda_N}{2}(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) + \lambda_N(||\hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1) \leq \frac{3}{2}\sqrt{k}\lambda_N||\hat{\nu}||_2$

As, $\frac{||X^Tw||_\infty}{N} \leq \frac{\lambda_N}{2}$ is assumed,

$$\frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) \leq \frac{\lambda_N}{2}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1)$$

$$\Rightarrow \frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) \leq \frac{\lambda_N}{2}(||\hat{\nu}_S||_1 + ||\hat{\nu}^c_S s||_1) + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) => \textbf{equation (4)}$$

$$\Rightarrow \frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) \leq \frac{3\lambda_N}{2}||\hat{\nu}_S||_1 - \frac{\lambda_N}{2}||\hat{\nu}_{S^c}||_1$$
$$\Rightarrow \frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) \leq \frac{3\lambda_N}{2}||\hat{\nu}_S||_1 => \textbf{equation (5)}$$

By L1-L2 norm inequality, $||x||_1 \leq \sqrt{n}||x||_2$,

$$||\hat{\nu}_S||_1 \leq \sqrt{k}||\hat{\nu}_S||_2$$

$$\Rightarrow \frac{3\lambda_N}{2}||\hat{\nu}_S||_1 \leq \frac{3}{2}\sqrt{k}\lambda_N||\hat{\nu}_S||_2 => \textbf{equation (6)}$$

Combining equations (4), (5) and (6),

$$\frac{\lambda_N}{2}(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) + \lambda_N(||\hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1) \leq \frac{3}{2}\sqrt{k}\lambda_N||\hat{\nu}||_2 => \textbf{equation (7)}$$

Combining equality from part (d) $\frac{||X\hat{\nu}||_2^2}{2N} \leq \frac{||X^Tw||_\infty}{N}||\hat{\nu}||_1 + \lambda_N(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1)$ and equation (7),

$$\frac{||\boldsymbol{X}\hat{\nu}||_2^2}{2N} \leq \frac{\lambda_N}{2}(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) + \lambda_N(||\hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1) \leq \frac{3}{2}\sqrt{k}\lambda_N||\hat{\nu}||_2$$

f)

To show, $||\hat{\beta} - \beta^*||_2 \leq \frac{3}{\gamma}\sqrt{\frac{k}{N}}\sqrt{N}\lambda_N$

Assuming Lemma 11.1 is true, the restricted eigenvalue condition from part (a)

$$\frac{\frac{1}{N}v^TX^TXv}{||v||_2^2} \geq \gamma \;\; \forall \text{ non-zero } v \in C$$

$$\Rightarrow \frac{||x\hat{\nu}||_2^2}{N||v||_2^2} \geq \gamma \;\; \forall \text{ non-zero } v \in C$$

$$\Rightarrow \frac{||x\hat{\nu}||_2^2}{2N} \geq \frac{\gamma||v||_2^2}{2} => \textbf{equation (8)}$$

Combining inequality $\frac{||\boldsymbol{X}\hat{\nu}||_2^2}{2N} \leq \frac{\lambda_N}{2}(||\hat{\nu}_S||_1 - ||\hat{\nu}_{S^c}||_1) + \lambda_N(||\hat{\nu}_S||_1 + ||\hat{\nu}_{S^c}||_1) \leq \frac{3}{2}\sqrt{k}\lambda_N||\hat{\nu}||_2$ and equation (8), we get

$$\frac{\gamma||\hat{\nu}||_2^2}{2} \leq \frac{3}{2}\sqrt{k}\lambda_N||\hat{\nu}||_2$$

$$\Rightarrow ||\hat{\nu}||_2 \leq \frac{3}{\gamma}\sqrt{k}\lambda_N$$

$$\Rightarrow ||\hat{\nu}||_2 \leq \frac{3}{\gamma}\sqrt{\frac{k}{N}}\sqrt{N}\lambda_N$$

As, $\hat{\nu} = \hat{\beta} - \beta^*$

$$||\hat{\beta} - \beta^*||_2 \leq \frac{3}{\gamma}\sqrt{\frac{k}{N}}\sqrt{N}\lambda_N$$

g)

The bound $\lambda_N \geq 2\frac{||X^Tw||_\infty}{N}$ shows up at two places in the proof.

1. From going to inequality 11.22 to inequality 11.23 on page 298, the bound $\lambda_N \geq 2\frac{||X^Tw||_\infty}{N}$ is assumed. In terms of the noise vector w, this simply informs us how to determine the regularization parameter λ_N.

2. Above inequality 11.44a on page 313, it is showed that if the bound $\lambda_N \geq 2\frac{||X^Tw||_\infty}{N}$ satisfies then the Lagrangian lasso error satisfies a bound that is it obeys the cone constraint.

h)

In order to have unique solution to our loss function $f_N(\beta) = \frac{1}{2N}||y - X\beta||_2^2$, $\beta\epsilon R^p$, it needs to be strongly convex. But unfortunately it is not strongly convex because $\nabla^2 f(\beta) = \frac{X^T X}{N}$, $\beta\epsilon R^p$, $X^T X$ has a rank of at most $min\{N, p\}$, so it is always rank deficient and hence not strongly convex as N is less than p. For this reason, we relax the strong convexity notion and make it necessary for some subset $C\epsilon R^p$. For appropriate choices of regularization parameter λ_N, β^* is supported by the subset $S = S(\beta^*)$ and the lasso error satisfies a cone constraint of the form $||\hat{\nu}_{S^c}||_1 \leq \alpha||\hat{\nu}_S||_1$ for some constant $\alpha \geq 1$. As a result, we limit our demand for strong convexity to vectors v that lie inside this cone.

i)

This theorem gives the following lasso error bound for Gaussian model error w,

$$||\hat{\beta} - \beta^*||_2 \leq \frac{c\sigma}{\gamma}\sqrt{\frac{\gamma klog\ p}{N}}$$

Whereas theorem 3 defined in class has the error that is upper bounded by $O(3\sigma\sqrt{N})$ as in the Gaussian distribution, with great probability the value lies between -3σ and 3σ. Hence, we can say that this theorem gives a better tight bound than theorem 3 for sparse signals.This theorem yields bounds that are minimax optimum, which means that no estimator can improve on them significantly. The disadvantage of this theorem over theorem 3 is that it only works for the signals which are exactly sparse and does not for the signals which are less sparse whereas theorem 3 can handle the less sparse signals.

j)

The common thread between the bounds on the 'Dantzig selector' and the LASSO is that they are approximately equivalent in terms of error bound and they both satisfy the cone constraint. If the noise vector is a zero mean Gaussian with standard deviation σ for both 'Dantzig selector' and the LASSO then error bounds of 'Dantzig selector' and the LASSO are upper bounded by $O(\sigma\sqrt{\frac{\gamma klog\ p}{N}})$.

Question 3

To run the code : "run Q3.m"

a)

Filtered back-projection using the RamLak filter

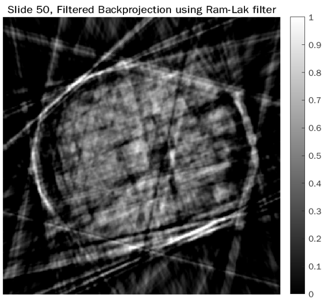

Figure 1: *
Slice 50

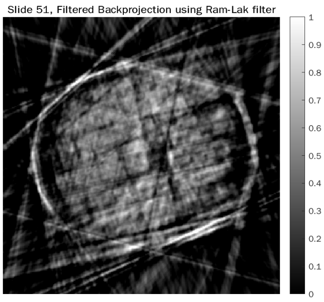

Figure 2: *
Slice 51

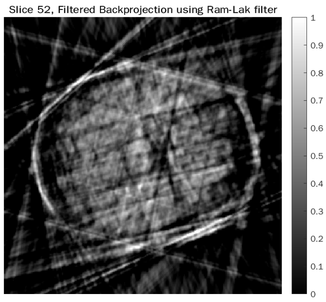

Figure 3: *
Slice 52

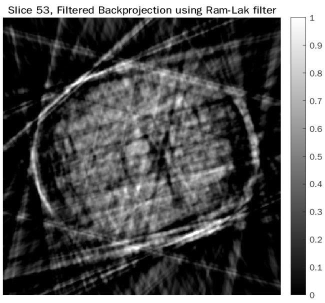

Figure 4: *
Slice 53

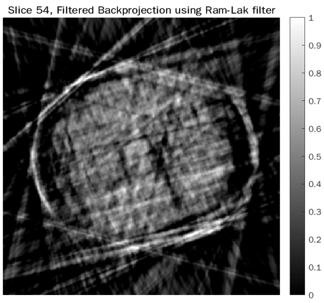

Figure 5: *
Slice 54

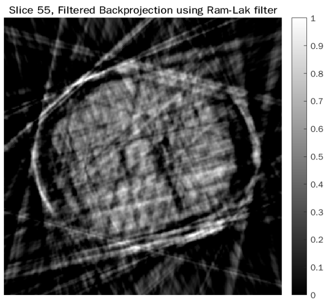

Figure 6: *
Slice 55

b)

Independent CS-based reconstruction

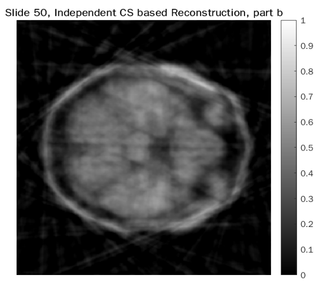

Figure 7: *
Slice 50

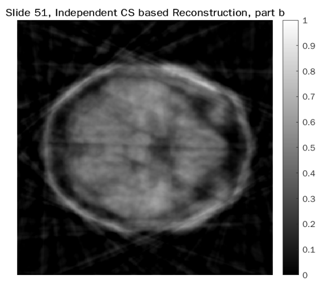

Figure 8: *
Slice 51

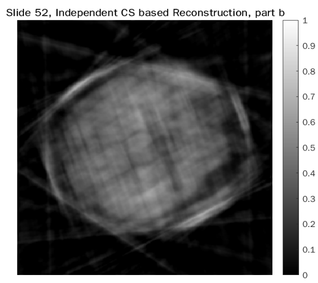

Figure 9: *
Slice 52

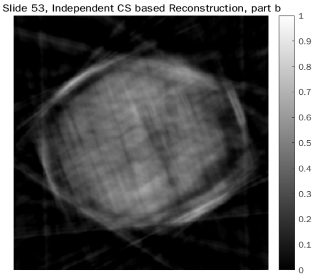

Figure 10: *
Slice 53

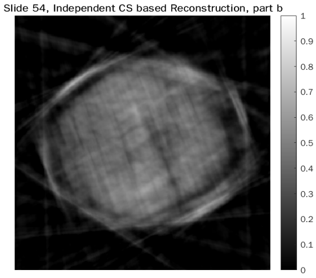

Figure 11: *
Slice 54

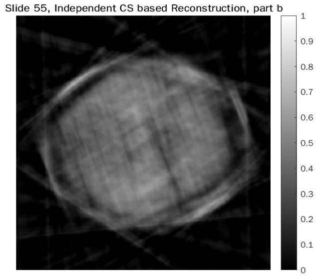

Figure 12: *
Slice 55

c)

Coupled CS-based reconstruction with two slices

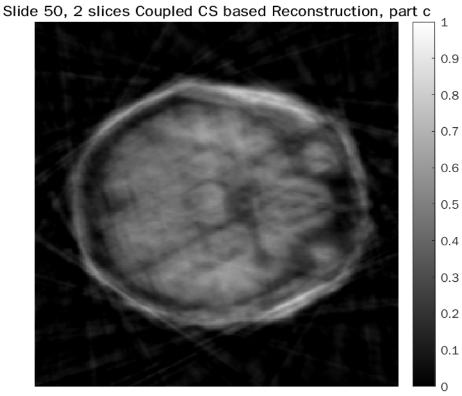

Figure 13: *
Slice 50

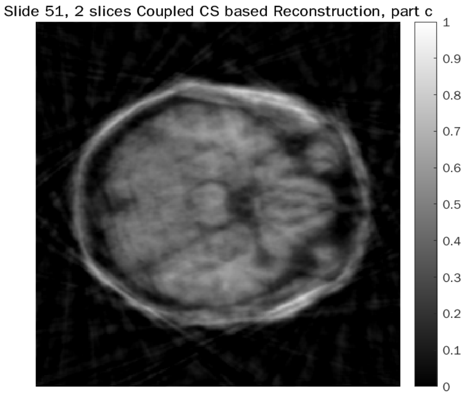

Figure 14: *
Slice 51

Coupled CS-based reconstruction with three slices

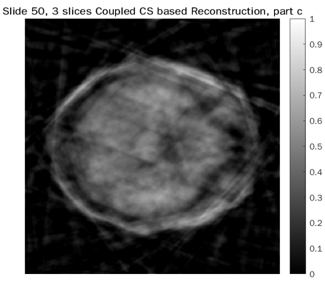

Figure 15: *
Slice 50

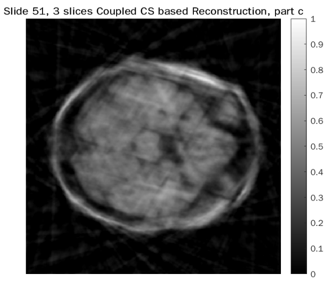

Figure 16: *
Slice 51

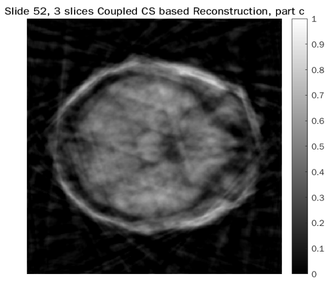

Figure 17: *
Slice 51

Question 4

- **Title of the Paper :** Big Data infrastructure for agricultural tomographic images reconstruction
- **Authors :** Gabriel M. Alves, Paulo E. Cruvinel
- **Published on :** 2018
- **Venue :** 12th International Conference on Semantic Computing, Laguna Hills, CA, USA

- **Link :** `https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8334494`

Mathematical problem defined in this paper

Considering the beam of instensity $I(x)$ travelled distance x through the object that has attenuation cofficient $\mu(x)$, then the total attenuation suffered by the beam is determined by the below equation,

$$I = I_0 exp(-\int_L \mu(x)dx)$$

By the means of inversion operation of above equation,

$$P(L) = \int_L \mu(x)dx = -ln(\frac{I}{I_0}) \quad => \ equation(1)$$

where P(L) is a projection of line L. In general, given the set of measures, P(L) of the object, we try to estimate and calculate the attenuation distribution µ(x).
By making use of dirac delta operator which has the sampling property and equation (1) in the two-dimensional case, can be written as,

$$P_\theta(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mu(x)\delta(xcos\theta + ysin\theta - t)dxdy$$

This above equation is Radon transform, $R_\theta\mu(t) = P_\theta(t)$
The problem of reconstructing an image consists of determining $\mu(x, y)$ from $P_\theta(t)$. The Radon Transform maps the spatial domain (x, y) into the domain (t, θ), where each point in space (t, θ) corresponds to a line in space (x, y).

Method of optimization that the paper uses to solve this problem

The Fourier slice theorem states that the Fourier transform of a projection of the 2D object along some direction θ (i.e. G(μ, θ)) is equal to a slice of the 2D Fourier transform of the object along the same direction θ (in the frequency plane), passing through the origin.

$$G(\mu, \theta) = [F(u, v)]_{u=\mu cos\theta, v=\mu sin\theta} = F(\mu cos\theta, \mu sin\theta)$$

Filtered Back Projection is used to reconstruct the image from a set set of projections based on Fourier slice theorem.

$$f(x, y) = \int_0^{2\pi}\int_0^{\infty} G(\mu, \theta)exp(j2\pi\mu(xcos\theta + ysin\theta))\mu d\mu d\theta$$

By applying a one-dimensional Fourier transform in the different projections at different angles we find an approximation of F(u, v) which is an approximation of (x, y) in the domain of frequency and reconstruction, so it is an Inverse Fourier Transform.

Question 5

Given, $R_\theta(f)$ is the radon transform of the image $f(x, y)$ in the given direction θ.
Let, $h(x, y) = f(\alpha x, \alpha y)$ where $\alpha \neq 0$

The radon transform is given as,

$$
\begin{aligned}
R_\theta(h) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\delta(xcos\theta + ysin\theta - \rho)dxdy \quad (putting\ x = \frac{x}{\alpha},\ y = \frac{y}{\alpha}) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\frac{x}{\alpha}, \frac{y}{\alpha})\delta(\frac{x}{\alpha}cos\theta + \frac{y}{\alpha}sin\theta - \rho)d(\frac{x}{\alpha})d(\frac{y}{\alpha}) \\
&= \frac{1}{\alpha^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(\frac{xcos\theta + ysin\theta - \rho\alpha}{\alpha})dxdy \\
&= \frac{1}{\alpha^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(xcos\theta + ysin\theta - \rho\alpha)dxdy \\
&= \frac{1}{\alpha^2}R_\theta(f)
\end{aligned}
$$

Hence, the radon transform of the scaled image $f(\alpha x, \alpha y)$ at $(\rho,\ \theta)$ where $\alpha \neq 0$ is $\frac{1}{\alpha^2}R_\theta(f)$ at $(\rho\alpha,\ \theta)$.

Question 6

The radon transform is given as,

$$
R(f)(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(xcos\theta + ysin\theta - \rho)dxdy
$$

Radon transform of unit impulse function $\delta(x,y)$

The radon transform of unit impulse function $\delta(x,y)$ is,

$$
R(\delta(x,y))(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x,y)\delta(xcos\theta + ysin\theta - \rho)dxdy
$$

As, $\delta(x,y)$ is zero at all (x,y) points except at $(x,y) = (0,0)$ the value of $\delta(x,y)$ is one,

$$
\begin{aligned}
R(\delta(x,y))(\rho,\theta) &= \delta(0,0)\delta(0+0-\rho) \\
&= \delta(0,0)\delta(-\rho) \\
&= \delta(-\rho) \qquad => equation(1)
\end{aligned}
$$

Radon transform of the shifted unit impulse $\delta(x - x_0, y - y_0)$

The radon transform of the shifted unit impulse $\delta(x - x_0, y - y_0)$ is,

$$
R(\delta(x - x_0, y - y_0))(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x - x_0, y - y_0)\delta(xcos\theta + ysin\theta - \rho)dxdy
$$

Putting $x = x + x_0\ and\ y = y + y_0$,

$$
\begin{aligned}
R(\delta(x - x_0, y - y_0))(\rho,\theta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x,y)\delta((x+x_0)cos\theta + (y+y_0)sin\theta - \rho)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x,y)\delta(xcos\theta + ysin\theta + x_0cos\theta + y_0sin\theta - \rho)dxdy \\
&= R(\delta(x,y))(-x_0cos\theta - y_0sin\theta + \rho, \theta) \quad (from\ equation(1)) \\
&= \delta(x_0cos\theta + y_0sin\theta - \rho)
\end{aligned}
$$