Rohan Chandra - hmwk5 Solutions

Question 1

a.) L2 Denoising Problem: min $\dfrac{\mu}{2}\|x\|^2 + \dfrac{1}{2}\|x-b\|^2$
The gradient of this would be $\mu\nabla^T\nabla x + (x-b) = 0$

This is equal to $\mu\nabla^T\nabla x + x = b$

Which can be rewritten in the form of $(\mu\nabla^T\nabla + I)x = b$, where $A = \mu\nabla^T\nabla + I$ and $\nabla^T\nabla$ is the laplacian kernel

0.1 Richardson

c.)

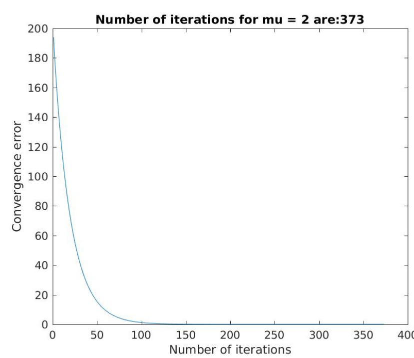

Figure 1: Convergence Plot for $\mu = 2$. Number of iterations are 373

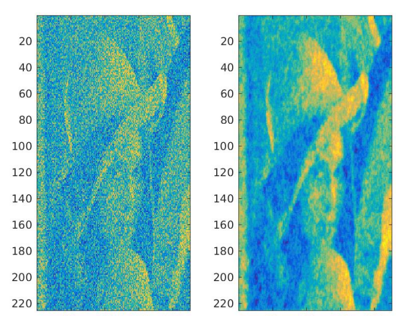

Figure 2: denoised image for $\mu = 2$

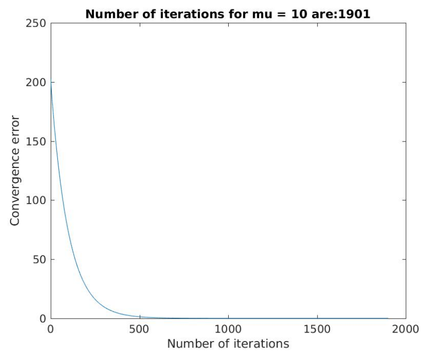

Figure 3: Convergence Plot for $\mu = 10$. Number of iterations are 1901

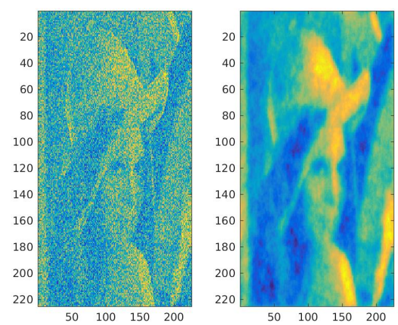

Figure 4: denoised image for $\mu = 10$

0.2 Conjgrad

d.)

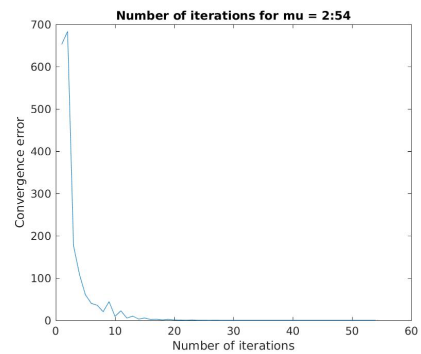

Figure 5: Convergence Plot for $\mu = 2$. Number of iterations are 54

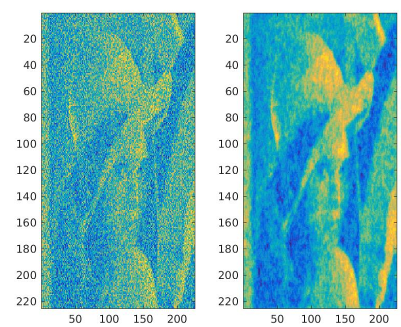

Figure 6: denoised image for $\mu = 2$

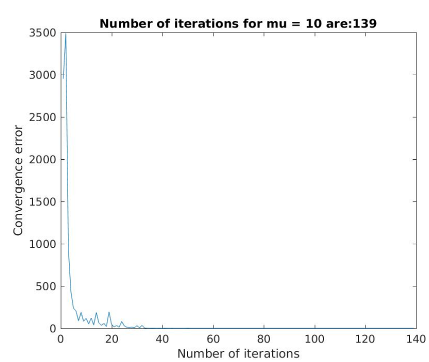

Figure 7: Convergence Plot for $\mu = 10$. Number of iterations are 139

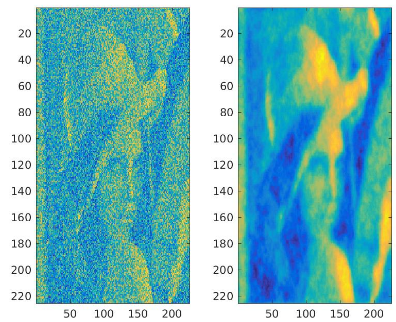

Figure 8: denoised image for $\mu = 10$

e.) It is observed that iteration counts increases both for Richardson iteration and conjugate gradient method.

Stability Restriction

According to the stability criterion, we can write out the taylor approximation of a convex function as, $f(y) \leq f(x) + (y-x)^T \nabla f(x) + \frac{M}{2}\|y-x\|^2$.

Now moving in a step size τ in the direction of the gradient, we can re-write the equation as

$$f(y - \tau \nabla f(x)) \leq f(x) - \tau \nabla f(x))^T \nabla f(x) + \frac{M\tau^2}{2}\|\nabla f(x)\|^2$$

This is equal to

$$f(y - \tau \nabla f(x)) \leq f(x) + \frac{M\tau^2 - 2\tau}{2}\|\nabla f(x)\|^2$$

$$\frac{M\tau^2 - 2\tau}{2} < 0 \implies \tau < \frac{2}{M}$$

This means that the the convergance rate is always less than $\frac{M}{2}$. Applying this to our problem, we can factorize our problem as

$$\frac{1}{2}\mu\|\nabla x\|^2 + \|x - b\|^2$$

.

This can be re-written as $\frac{1}{2}\mu x^T\nabla^T\nabla x + x^Tx - 2x^Tb + \|b\|^2$. This gives

$$\frac{\mu}{2}x^T(\nabla^T\nabla + I)x - 2x^Tb + \|b\|^2$$

$\dfrac{1}{2}x^T(\mu\nabla^T\nabla + \mu I)x - 2x^Tb + \|b\|^2$ is a quadratic where $\nabla^T\nabla$ is the laplacian and $\mu\nabla^T\nabla + \mu I$ is the hessian matrix. Let λ be the largest eigenvalue of this hessian. That would be the largest curvature of the quadratic and convergence is defined when

$$\tau < \frac{2}{\lambda}$$

By this, we can see that τ depends inversely on μ as $\dfrac{2}{largest\quad eigenvalue\quad of(\mu\nabla^T\nabla + \mu I)}$ Hence as μ increases, τ decreases and hence results in slower convergence.

0.3 L2Denoise

f.)

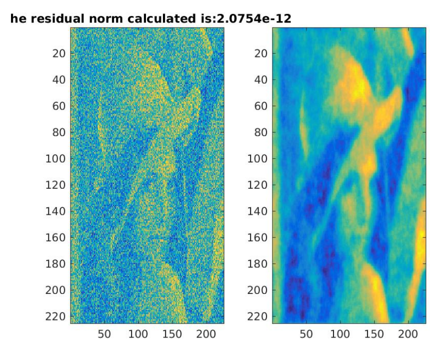

Figure 9: denoised image with $norm = e^{-12}$

Question 2

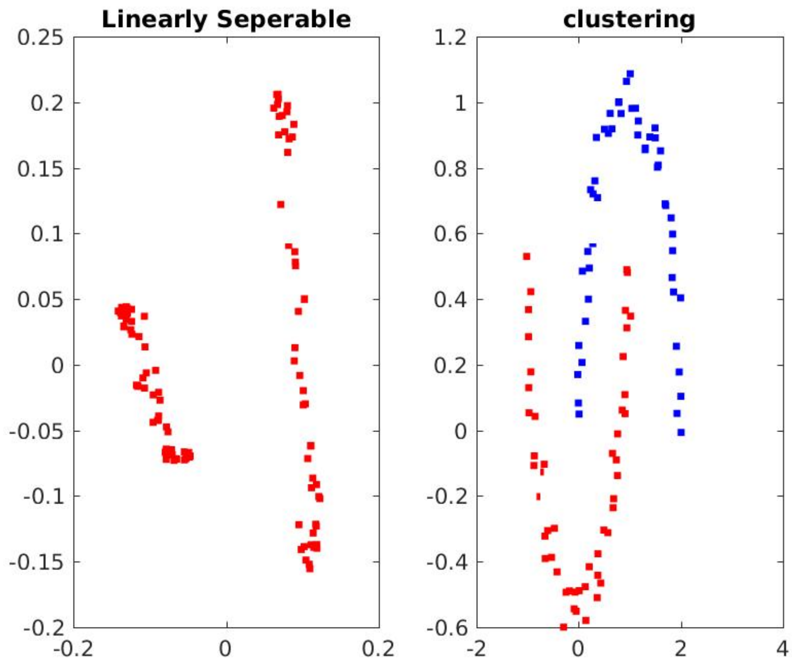

Figure 10: linear seperation and clustering for 100 points

Question 3

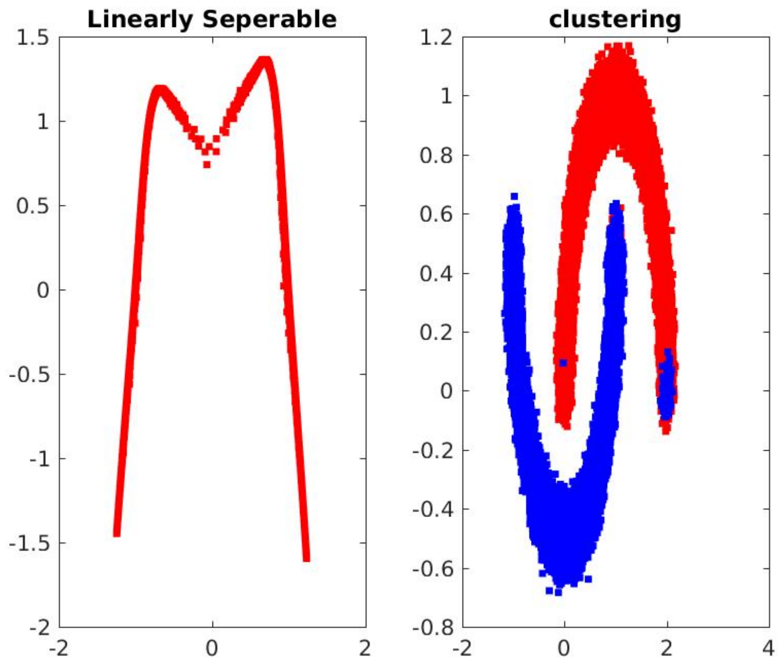

Figure 11: linear seperation and clustering for 100,000 points