CS 754 - Assignment 5

Nihar Mehta, Ushasi Chaudhuri

April 20, 2018

Question 1

Codes are attached. Please refer to Fig.2 and 2 for the DCT coefficient plot.

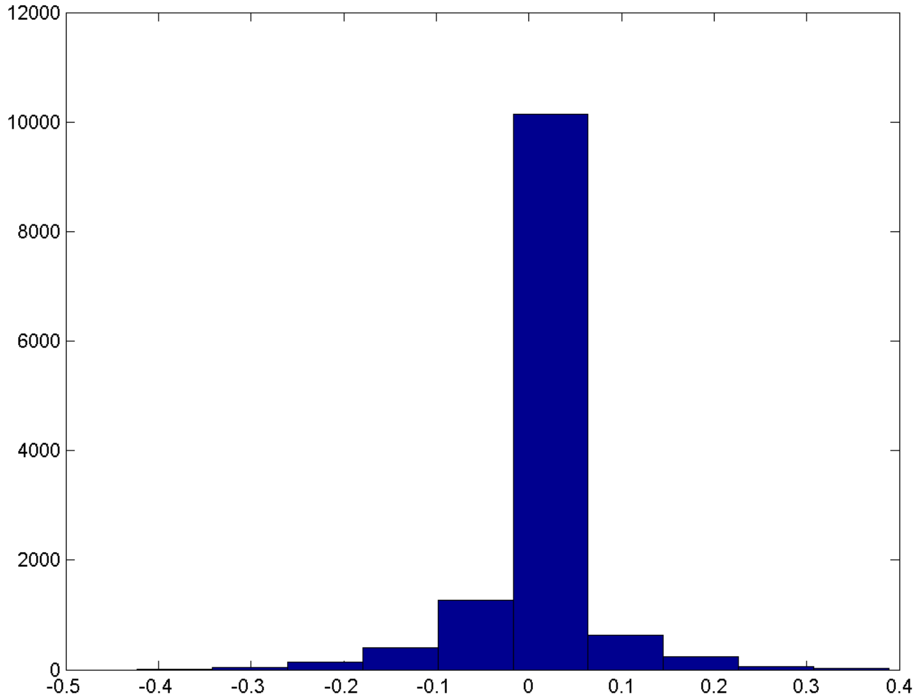


Figure 1: DCT Histogram of D5.

+

Question 2

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4288165
(i), (ii) Please find the attached images from Fig. 3 to 4 for the solution of part (i) and (ii). The solution to this question has been included in

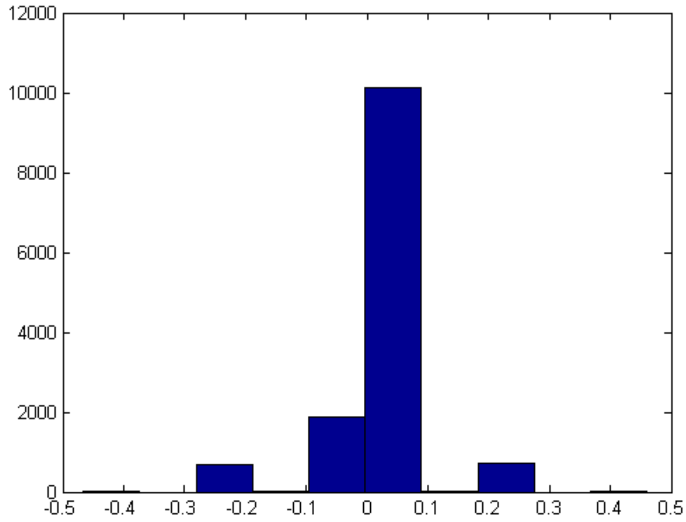


Figure 2: DCT Histogram of D6.

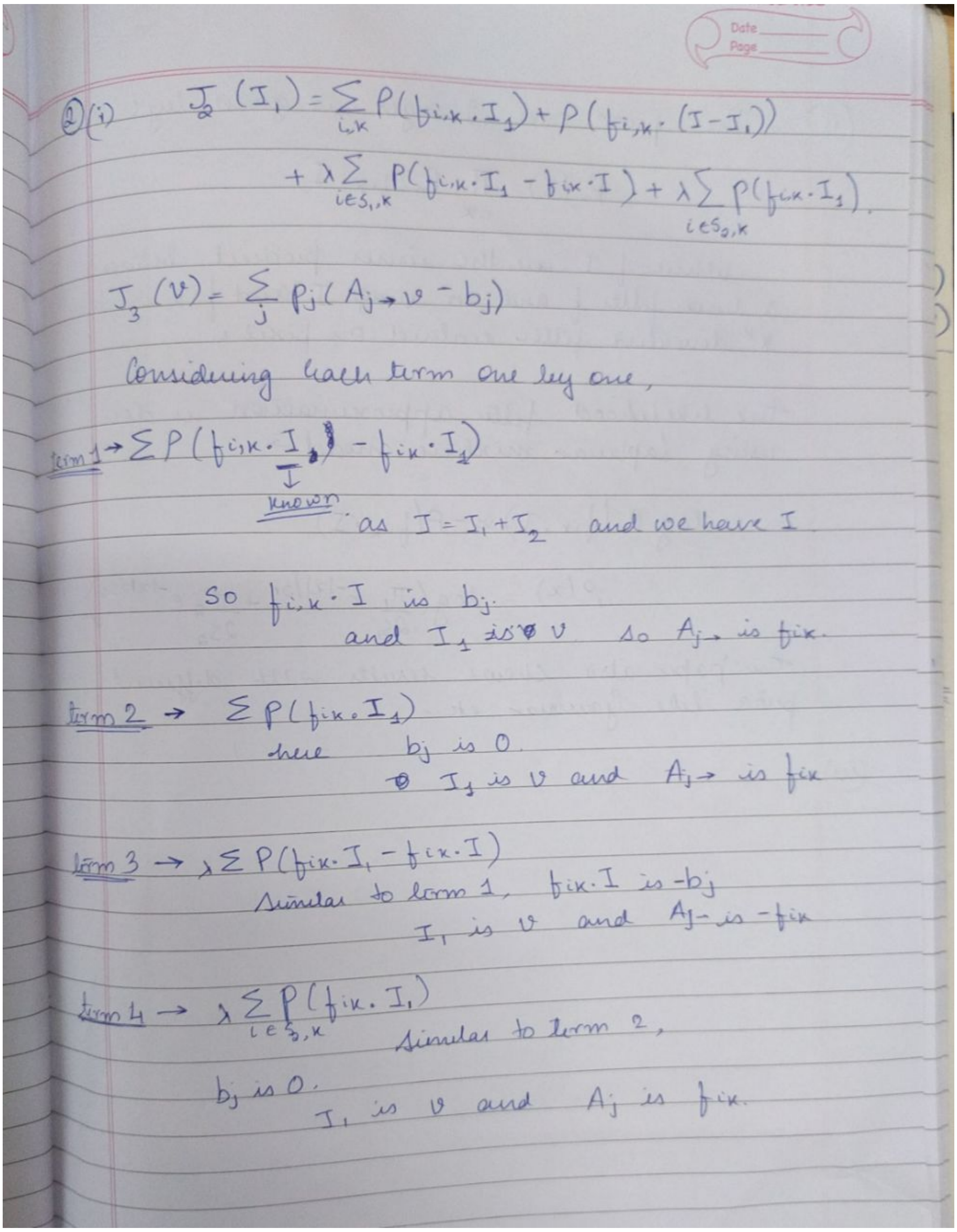
(iii) When we look at the logarithm of the histogram, the curve is always below the straight line connecting the maximum and minimum values (fig 2 in the given in paper 2). In order to contrast with the Gaussian distribution (that is always above the straight line) or the Laplacian distribution (that is simply a straight line in the log domain). The log distribution is always below the straight line. This is crucial for obtaining transparency decompositions from a single image. Distributions that are above the straight line will prefer to split an edge of unit contrast into two edges (one in each layer) with half the contrast, while distributions below the line will prefer decompositions in which the edge only appears in one of the layers but not in the other.

Although the Laplacian distributions are not sparse, the mixture is. The Laplacian distribution is exactly at the border between sparse and non sparse distributions. The Gaussian distribution is not sparse (it is always above the straight line) and distributions for which $\alpha \leq 1$ are sparse. In particular, a Laplacian distribution has been experimentally shown to be a very good fit for all except the DC coefficient (u = v = 0).

Question 3

Complex NMF: A New Sparse Representation For Acoustic Signals
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4960364
The following paper is in the domain of Audio-signal processing. This paper presents a sparse representation for acoustic signals which is based on a mixing model defined in the complex-spectrum domain (where additivity holds). The authors extract recurrent patterns of magnitude spectra that underlie complex spectra and the phase estimates of constituent signals. The authors derive an efficient iterative algorithm which reduces to the multiplicative update algorithm for non-negative matrix factorization.

NMF is a powerful tool for extracting regularities or structural patterns from acoustic signals. Because of the phase-invariant nature of magnitude spectra, NMF is able to project all signals that have the same spectral shape onto a single basis. This allows the representation of a variety of acoustic phenomena efficiently using a very compact set of spectrum bases. Complex NMF is based on mixing model defined in the complex-spectrum domain (where additivity holds). It can extract the recurrent patterns of magnitude spectra that underlie the observed complex spectra and the phase spectra of constituent signals,. Also, it can be performed with an efficient iterative algorithm, which reduces to the multiplicative update of the algorithm.



2 (i) $J_2(I_1) = \sum_{i,k} P(f_{i,k} \cdot I_1) + P(f_{i,k} \cdot (I - I_1))$

$+ \lambda \sum_{i \in S_{1,k}} P(f_{i,k} \cdot I_1 - f_{ik} \cdot I) + \lambda \sum_{i \in S_{0,k}} P(f_{ik} \cdot I_1)$.

$J_3(v) = \sum_j \rho_j(A_{j\to} v - b_j)$

Considering each term one by one,

term 1 → $\sum P(f_{i,k} \cdot \underbrace{I}_{\text{known}} - f_{ik} \cdot I_1)$

as $I = I_1 + I_2$ and we have I

so $f_{i,k} \cdot I$ is b_j.

and I_1 is v so $A_{j\to}$ is f_{ik}.

term 2 → $\sum \rho(f_{ik} \cdot I_1)$

here b_j is 0.

I_1 is v and $A_{j\to}$ is f_{ik}

term 3 → $\lambda \sum P(f_{ik} \cdot I_1 - f_{ik} \cdot I)$

Similar to term 1, $f_{ik} \cdot I$ is $-b_j$

I_1 is v and $A_{j\to}$ is $-f_{ik}$

term 4 → $\lambda \sum_{i \in S_{0,k}} P(f_{ik} \cdot I_1)$

Similar to term 2,

b_j is 0.

I_1 is v and A_j is f_{ik}.

Figure 3: Question.2 (part i).

Short-term Fourier transform of an arbitrary acoustic signal $F_{x,t} \in C$. It consists of k complex valued elements.

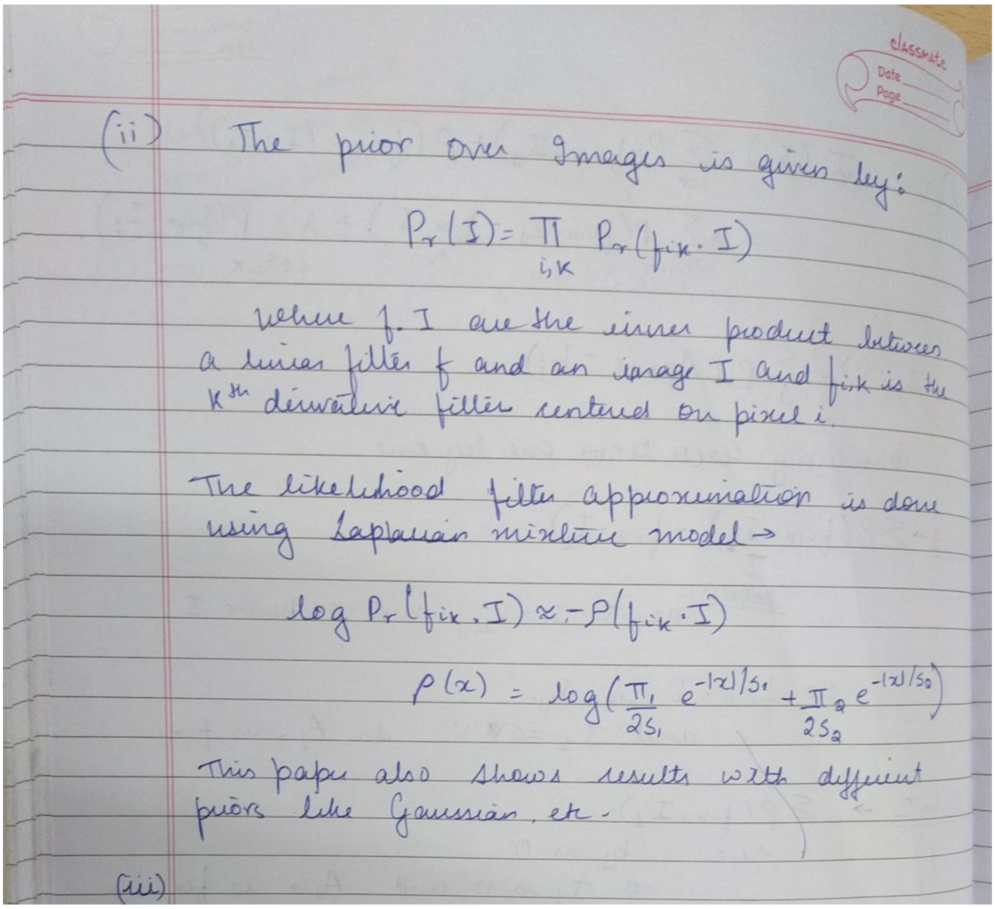
$$F_{x,t} = \sum_{k=1}^{K} |a_{k,x,t}| e^{j\phi_{k,x,t}} \tag{1}$$

where x and t are the frequency and frame indices. Factorize $|a_{k,x,t}|$ into the product of nonnegative parameters, $H_{k,x}$ and $U_{k,t}$

$$|a_{k,x,t}| = H_{k,x} U_{k,t} \tag{2}$$

also,

$$H_{k,x} = 1 (k = 1, ...K) \tag{3}$$



(ii) The prior over Images is given by:

$$P_r(I) = \prod_{i,k} P_r(f_{ik} \cdot I)$$

where f.I are the inner product between a linear filter f and an image I and f_{ik} is the k^{th} derivative filter centered on pixel i.

The likelihood filter approximation is done using Laplacian mixture model →

$$\log P_r(f_{ik} \cdot I) \approx -\rho(f_{ik} \cdot I)$$

$$\rho(x) = \log\left(\frac{\pi_1}{2s_1} e^{-|x|/s_1} + \frac{\pi_2}{2s_2} e^{-|x|/s_2}\right)$$

This paper also shows results with different priors like Gaussian, etc.

(iii)

Figure 4: Question.2 (part ii).

In order to avoid an indeterminacy in the scaling, the following mixing model is given:

$$F_{x,t} = \sum^{K} H_{k,x} U_{k,t} e^{j\phi_{k,x,t}} \tag{4}$$

The optimization problem addressed here is as follows: minimize

$$f(\theta) = \sum_{x,t} |Y_{x,t} - F_{x,t}|^2 + 2\lambda \sum |U_{k,t}|^p \tag{5}$$

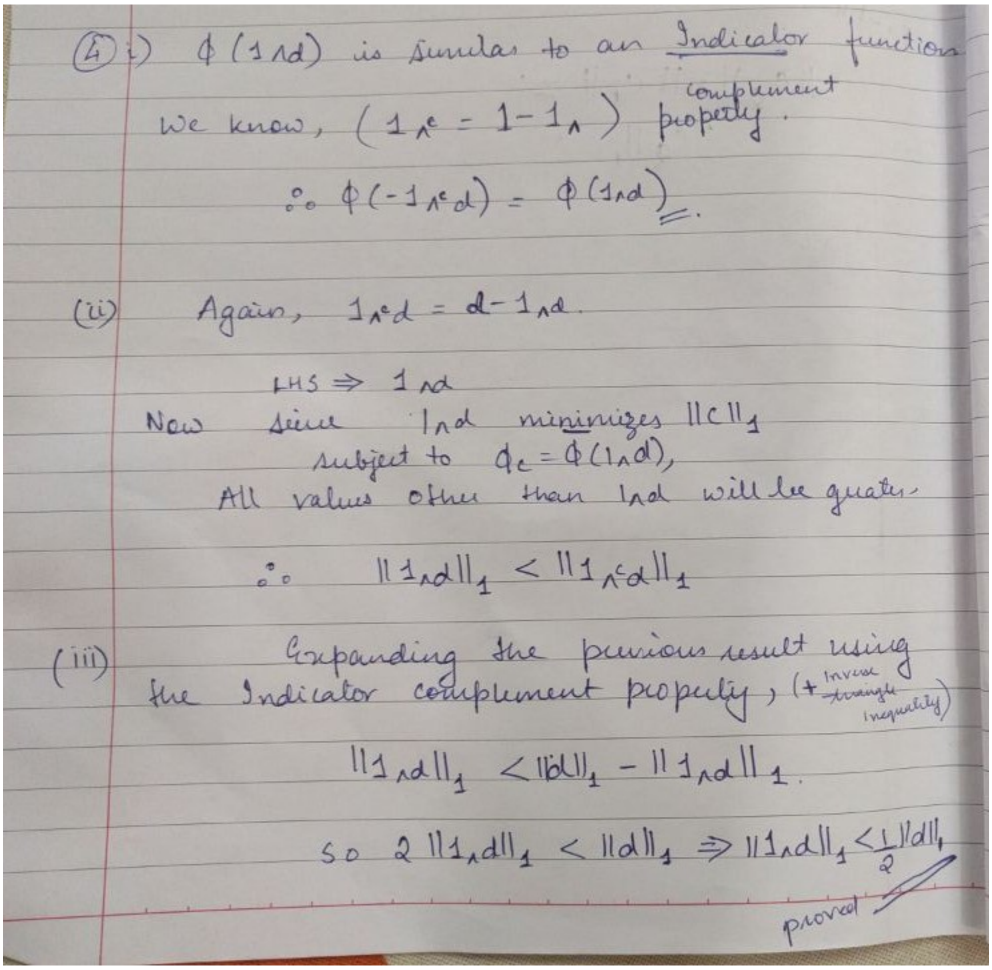
subject to

$$H_{k,x} = 1 (k = 1, ...K) \tag{6}$$

The authors use some of the speech-data for experimentations. The aim of the experiments was to determine whether Complex NMF has an effect as with NMF to extract the underlying patterns of magnitude spectra from audio data. For the first experiment a Complex NMF was tested for a singer speaker. For the second experiment the authors tested Complex NMF on a mixed voice signal. If the synthesized signal contained two voices, it suggests that the model is over-fitting their observations. If the synthesized signal does not sound like speech, it suggests that the model is under-fitting their observation.

Question 4

The solution to this question has been included in Fig. 5 to 8.



(4) i) $\phi(1_\Lambda d)$ is similar to an Indicator function

We know, $(1_{\Lambda^c} = 1 - 1_\Lambda)$ complement property.

$\therefore \phi(-1_{\Lambda^c} d) = \phi(1_\Lambda d)$

(ii) Again, $1_{\Lambda^c} d = d - 1_\Lambda d$.

LHS $\Rightarrow 1_\Lambda d$

Now since $1_\Lambda d$ minimizes $||c||_1$ subject to $\phi_c = \phi(1_\Lambda d)$,

All values other than $1_\Lambda d$ will be greater.

$\therefore ||1_\Lambda d||_1 < ||1_{\Lambda^c} d||_1$

(iii) Expanding the previous result using the Indicator complement property, (+ Inverse triangle inequality)

$||1_\Lambda d||_1 < ||d||_1 - ||1_\Lambda d||_1$.

So $2||1_\Lambda d||_1 < ||d||_1 \Rightarrow ||1_\Lambda d||_1 < \frac{1}{2}||d||_1$

proved

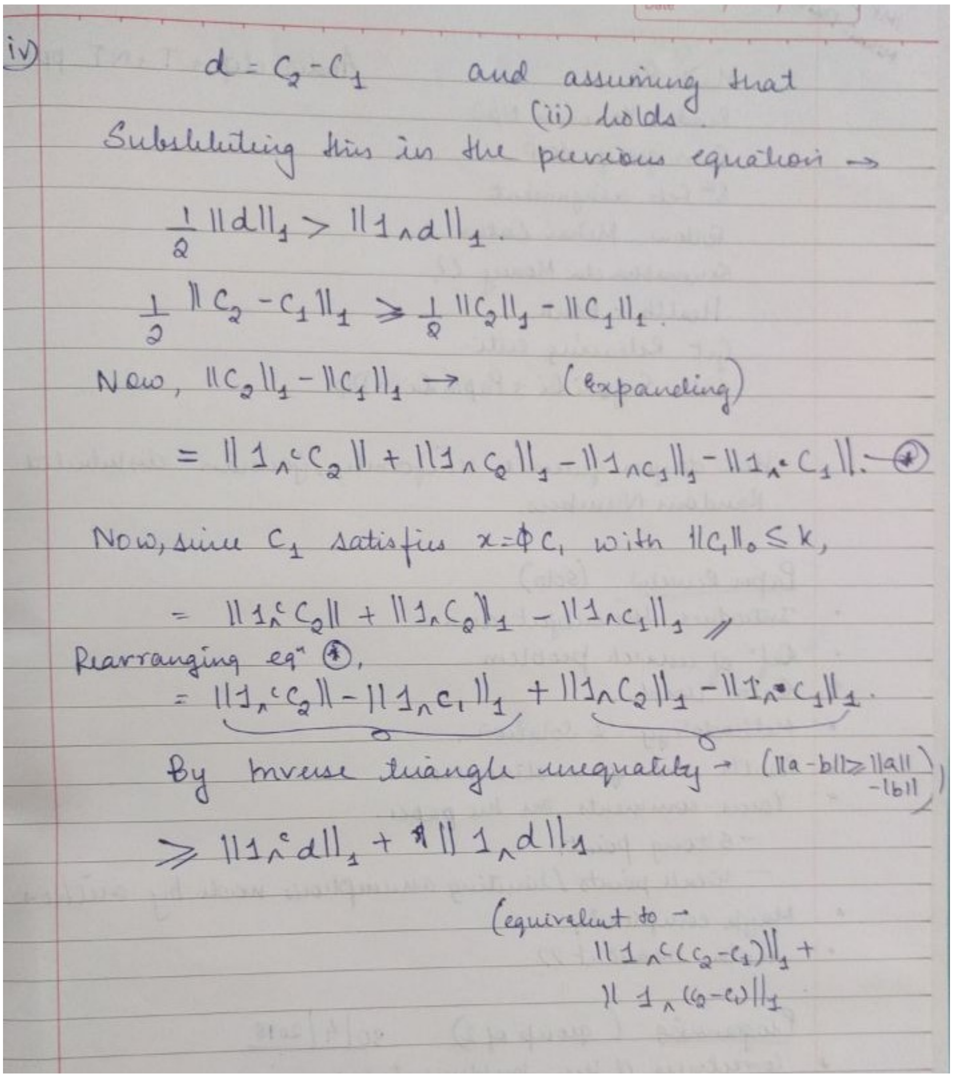
Figure 5: Question.4 (part 1).

Question 5

a)MAP estimate

x is taken from a gaussian distrubution with mean $0_{k\times 1}$ and covariance matrix Σ_x.
Hence the prior on x will be $p(x) = \frac{e^{-\frac{1}{2}x^T\Sigma^{-1}x}}{\sqrt{(2\pi)^k|\Sigma|}}$
The likelihood shall be $p(y|x) = e^{-\frac{1}{2\sigma^2}||y-\phi x||^2}$
The MAP estimate is given as:

$\hat{x}_{MAP} = argmax_x p(x|y) = argmax_x p(y|x)p(x)$
$= argmin_x \frac{1}{2\sigma^2}||y - \phi x||^2 + \frac{1}{2}x^T\Sigma^{-1}x$
Differentiating with respect to x and equating it to 0. Thus we shall have
$\frac{1}{\sigma^2}(-\phi^T y + \phi^T\phi\hat{x}_{MAP}) + \Sigma^{-1}\hat{x}_{MAP} = 0$
$\hat{x}_{MAP} = (\phi^T\phi + \sigma^2 * \Sigma^{-1})^{-1}(\phi^T y)$



iv) $d = c_2 - c_1$ and assuming that (ii) holds.

Substituting this in the previous equation →

$$\frac{1}{2}\|d\|_1 > \|1_\wedge d\|_1.$$

$$\frac{1}{2}\|c_2 - c_1\|_1 \geq \frac{1}{2}\|c_2\|_1 - \|c_1\|_1.$$

Now, $\|c_2\|_1 - \|c_1\|_1 \rightarrow$ (expanding)

$$= \|1_\wedge^c c_2\| + \|1_\wedge c_2\|_1 - \|1_\wedge c_1\|_1 - \|1_\wedge^c c_1\|. \quad — (*)$$

Now, since c_1 satisfies $x = \Phi c_1$ with $\|c_1\|_0 \leq k$,

$$= \|1_\wedge^c c_2\| + \|1_\wedge c_2\|_1 - \|1_\wedge c_1\|_1$$

Rearranging eq^n (*),

$$= \underbrace{\|1_\wedge^c c_2\| - \|1_\wedge c_1\|_1}_{} + \underbrace{\|1_\wedge c_2\|_1 - \|1_\wedge^c c_1\|_1}_{}.$$

By inverse triangle inequality → ($\|a-b\| \geq \|a\| - \|b\|$)

$$\geq \|1_\wedge^c d\|_1 + \|1_\wedge d\|_1$$

(equivalent to → $\|1_\wedge^c (c_2 - c_1)\|_1 + \|1_\wedge (c_2 - c_1)\|_1$)

Figure 6: Question.4 (part 2).

The graphs are attached:

Date / /

(vi) $\Phi d = 0$

$\Phi * \Phi d = 0$

Now, $*$ is a hodge-star operator

so $\Phi * \Phi d \Rightarrow \langle \Phi_i, \Phi_j \rangle d$.

~~Now, considering this elementwise,~~ we get $\rightarrow$

Now, $\Phi = (\varphi_i)_{i \in I}$.

so $\langle \Phi_i, \Phi_j \rangle d \rightarrow$ can be converted as

$\langle \varphi_i, \varphi_j \rangle d$ (Note $\rightarrow d = d_i + d_j$

Considering this element wise,

$$d_i + \sum_{i \neq j} \langle \varphi_i, \varphi_j \rangle d_j = 0.$$

$$\text{so } d_i = -\sum_{i \neq j} \langle \varphi_i, \varphi_j \rangle d_j$$

(vii) from the equation $\rightarrow$

$$|d_i| \leq \left(1 + \frac{1}{\mu(\Phi)}\right)^{-1} \|d\|_1.$$

Now, $1_\Lambda d$ is the unique minimizer of $\|C\|_1$, subject to $\Phi c = \Phi(1_\Lambda d)$

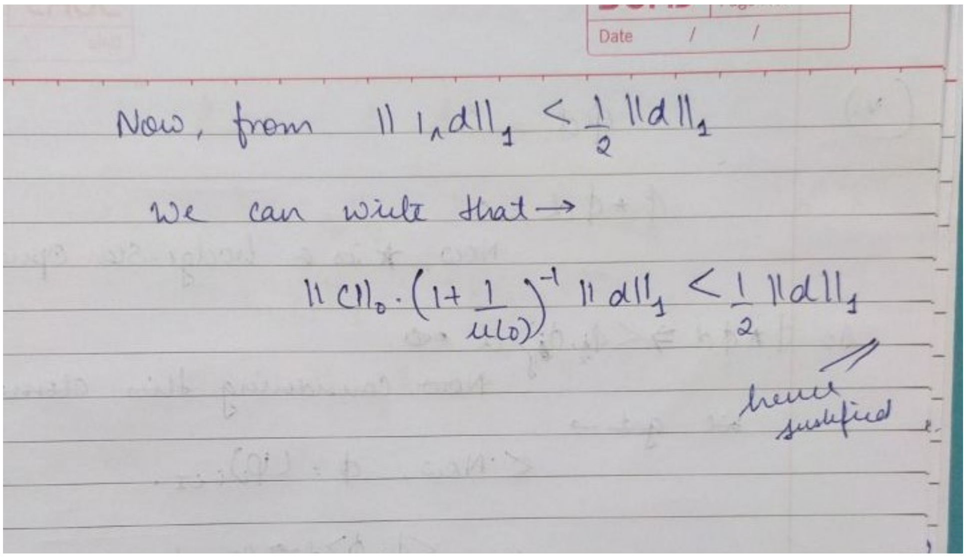
but here, $\Phi d = 0$. $|\Lambda| = \|C\|_0$ so, $\Phi \Lambda = \Phi(1_\Lambda d)$

LHS $\cdot$ so $\|1_\Lambda d\|_1 \leq |\Lambda| \cdot \left(1 + \frac{1}{\mu(\Phi)}\right)^{-1} \|d\|_1$

Substituting $\|C\|_0$ (since $|\Lambda| = \|C\|_0$)

we get, $= \|C\|_0 \left(1 + \frac{1}{\mu(\Phi)}\right)^{-1} \|d\|_1$

Figure 7: Question.4 (part 3).



Date / /

Now, from $\| 1_\Lambda d \|_1 < \frac{1}{2} \|d\|_1$

we can write that →

$$\|C\|_0 \cdot \left(1 + \frac{1}{\mu(0)}\right)^{-1} \|d\|_1 < \frac{1}{2} \|d\|_1$$

hence justified

Figure 8: Question.4 (part 4).

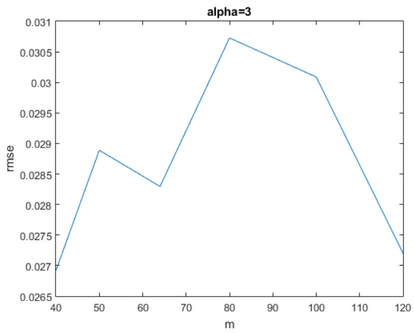


Figure 9: Question.5 (part i).

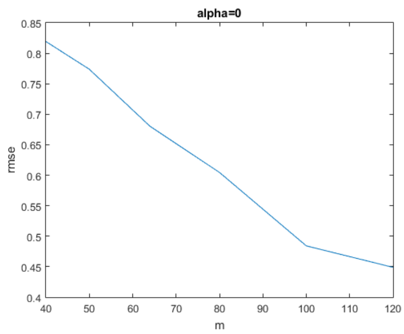


Figure 10: Question.5 (part ii).