Ed Qu

Homework 3

(Dated: March 2, 2016)

PREPROCESSING

According to the requirements in this homework, I defined a $g1_{Poisson}$ as the noisy data for this question. The function "poissrnd()" in Matlab generated Poisson random variables for the components of $g1_{Poisson}$. The mean of the Poisson random variable of each component was equal to the corresponding component of the noiseless data vector. In practice, the function "poissrnd()" output only integers, but the components in the noiseless $g1$ had a order of magnitude about 10^{-3} so that the statement, poissrnd($g1$), would generate a long binary vector. The information necessary for image reconstruction was overwhelm by the Poisson noise.

To add the Poisson noise and preserve the meaning of the measured data, I multiplied the vector $g1$ by a factor 10^4 and by a factor of 10^7, respectively. Using the products as the input, "poissrnd()" generated two measured data, $g1_{p1}$ and $g1_{p2}$. $g1_{p1}$ had a higher noise level than $g1_{p2}$.

In contrast, the Gaussian noise was defined with respect to the maximum of $g1$. Multiplying a factor with $g1$ would not affect the final result.

QUESTION 1

The equation that I used in the EM algorithm was

$$f_j^{(k+1)} = \frac{f_j^{(k)}}{H_{+,j}} \sum_{i=1}^{M} H_{ij} \frac{g_i}{[Hf^{(k)}]_i} \tag{1}$$

here the vector f represented the reconstructed image; $f_j^{(k)}$ was the jth entry in the vector f in the kth iteration step. H was the operator describing the imaging system. H_{ij} was the entry of H; $H_{+,j}$ was the sum of H_{ij} over the jth column. $[Hf^k]_i$ was the ith row of the product Hf. g was the measured data; it could be either the noiseless $g1$ or the noisy $g1_{Poisson}$.

I show some reconstructed images in the Fig. 1. The iteration step at which the image was reconstructed is shown on the top. The images in the first, second, and third row corresponds to the noiseless $g1$, noisy $g1_{p1}$, and noisy $g1_{p2}$, respectively. Here, we can observe a clear decrease of the noise level in the first row, as the iteration steps increases. After 500 iteration steps, we can recover most of important features in the phantom. In contrast, when the measured data $g1_{p1}$ includes a higher level Poisson noise, the EM algorithm does not converge to a meaningful result. The images in the second row at the 10th and 200th iteration show some features of the phantom, that is, $f^{(k)}$ is close to the true result f in the Hilbert space. Further study (not shown here) shows that the result $f^{(k)}$ reconstructed by $g1_{p1}$ will diverge after thousands of steps. When we decrease the noise level in the measured data g to an acceptable degree (the third row in the Fig. 1), the EM algorithm recover its ability of image reconstruction, and minimize the noise level in the image as the iteration step increases. We can recover the phantom once again after 500 iteration steps.

Fig. 2 shows more details about the convergence of the EM algorithm in the first 500 iteration steps. I show the variation of three criteria, root-mean-square error (RMSE), Kullback-Leibler (KL) distance, and least square $||g - Hf^{(k)}||$. The expression of RMSE and KL distance are

$$RMSE = \sqrt{f^{(k)} - f} \tag{2}$$

$$KL = \sum_{i=1}^{M} g_i \ln\left[\frac{g_i}{(Hf^{(k)})_i}\right] + (Hf^{(k)})_i - g_i \tag{3}$$

The results are consistent with the theory of EM algorithm. EM algorithm converges to a vector f_{EM} at which the KL distance is minimized. This fact will not change with the noise level in the measured data g (Fig. 2(b)), but the noise level indeed affect the rate of convergence. The RMSE shows a high sensitivity to the noise level (Fig. 2(b)). When the noise level surpass a critical point, for instance $g1_{p1}$, the noisy data and the noiseless data will converge to totally different results. The least square in Fig. 2(c) shows a similar result as the KL distance.

QUESTION 2

In this problem, I used the simultaneous SART algorithm to calculate the least square solution The equation that I used was

$$f_j^{(k+1)} = f_j^{(k)} + \frac{1}{H_{+,j}} \sum_{i=1}^{M} H_{ij} \frac{g_i}{[Hf^k]_i} \tag{4}$$

The definition of the variables in (4) was same as (1), and $H_{i,+}$ was the sum of H_{ij} over the ith row.

Fig. 3 shows some reconstructed image. For the noiseless $g1$, the noise level in the images decreases as the iteration step increase. The rate of convergence of the SART algorithm for $g1$ is much slower than that of EM method. As a result, I cannot recover some fine structures in the phantom at 500th iteration step. If a 3%

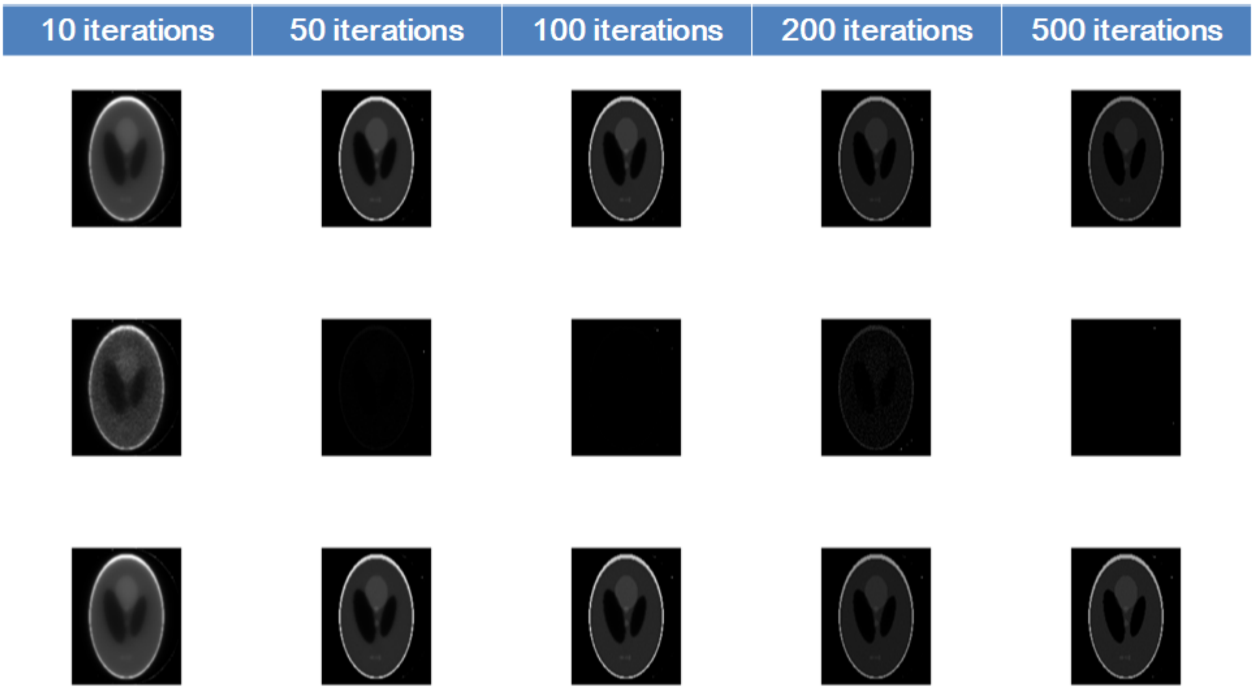

FIG. 1: Reconstructed images using EM algorithm. The image reconstructions used noiseless $g1$ (first row), noisy $g1_{p1}$ (second row), and noisy $g1_{p2}$ (third row), respectively.

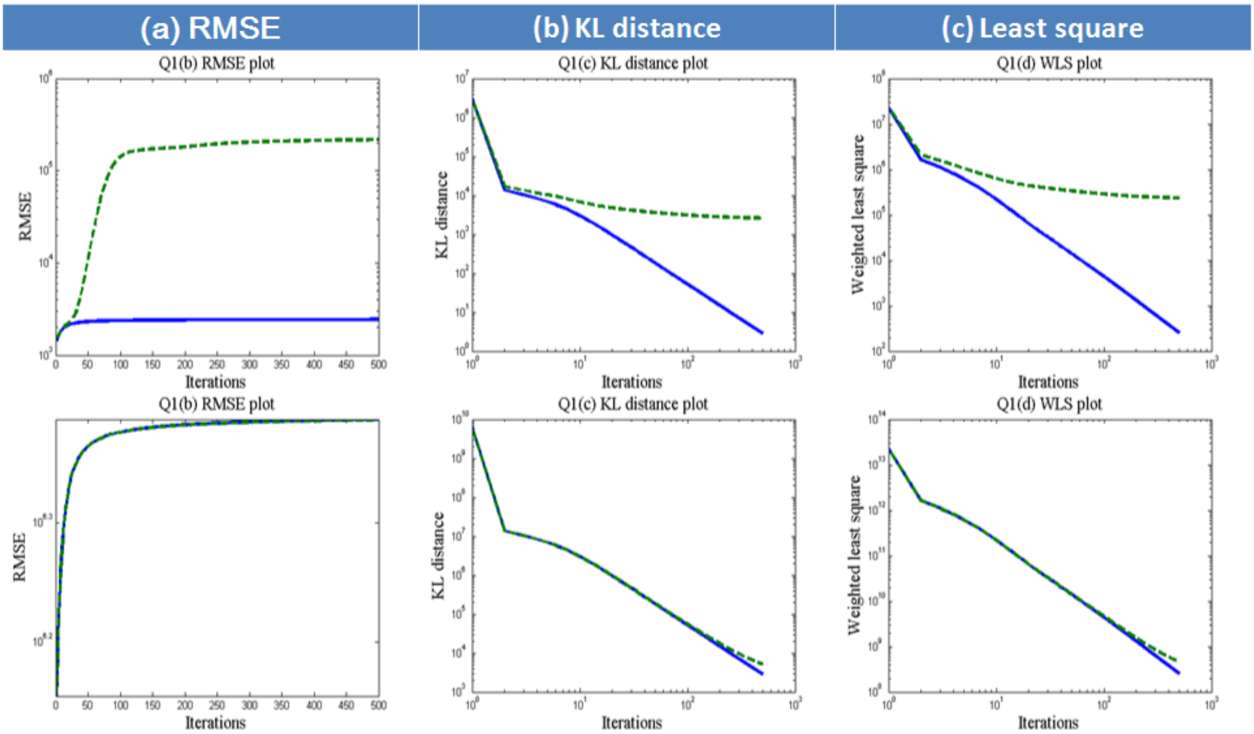

FIG. 2: The convergence of the EM algorithm. (a) RMSE plot. (b) KL distance plot. (c) Least square plot. Solid line: noiseless $g1$. Dash line: noisy $g1_{p1}$ (top row) or $g1_{p2}$ (bottom row).

FIG. 3: Reconstructed images using SART algorithm. The image reconstructions used noiseless $g1$ (top row) and noisy $g1$ (bottom row), respectively.

Gaussian white noise is added to the measure data $g1$, the SART algorithm cannot converge to the true result f anymore. Thus the reconstructed image from the noisy $g1$ at 500th iteration step shows nothing.

The plot of RMSE (Fig. 4) also support this conclusion.The RMSEs of the noiseless $g1$ and of noisy $g1$ are approaching to different values as the iteration steps increases. The SART algorithm minimizes the KL distance and the least square regardless of the noise level in the measured data $g1$, but its rate of convergence is affect by this noise.

QUESTION 3

EM algorithm can be expressed as

$$f_{MAP} = \arg\min_f \{-\operatorname{pr}(f|g)\} \tag{5}$$

$$= \arg\min_f \{-\ln \operatorname{pr}(f|g)\} \tag{6}$$

In practice, we will further decompose the (6) based on Bayes' theorem into

$$f_{MAP} = \arg\min_f \{-\ln[\operatorname{pr}(g|f)] - \ln[\operatorname{pr}(f)]\} \tag{7}$$

The first term in (7) on the right hand side is the likelihood function; the second term describe the probability distribution of the model parameters. If the measured data g are described by Poisson noise, we can neglect the second term and directly minimize the likelihood function. If the measured data g are described by Gaussian white noise, we must minimize the likelihood function and the second term simultaneously. In this case, more computational resources are required.

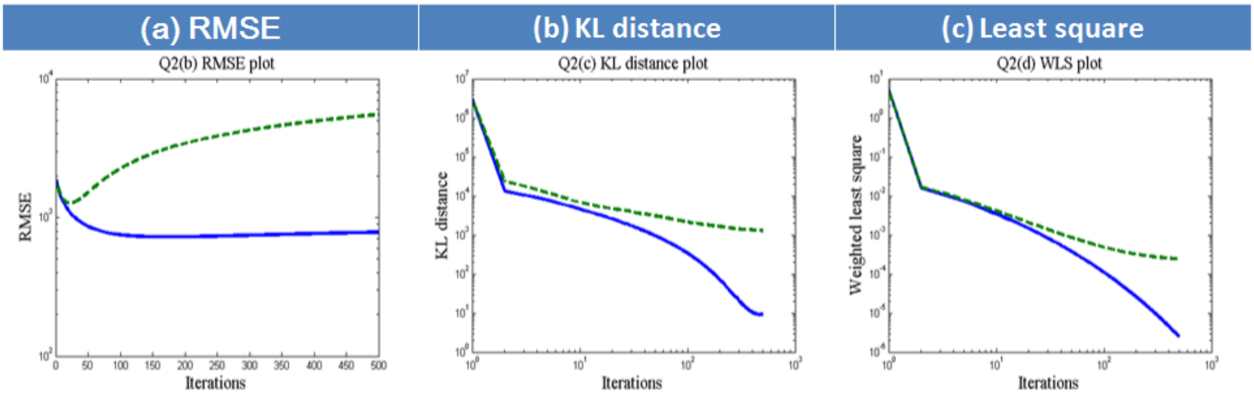

FIG. 4: The convergence of the SART algorithm. (a) RMSE plot. (b) KL distance plot. (c) Least square plot. Solid line: noiseless $g1$. Dash line: noisy $g1$.