Matlab Assessment

Please complete the Excel spreadsheet *Matlab Assessment* based on the following exercises.

Exercise 1

Run Program\_01 and find the following information for the file *MRI\_BREAST.* Ctrl-F (which will give you a find box) is useful here, but remember that the variables you are searching for aren’t necessarily labelled exactly as below.

Magnetic Field Strength

Echo Time

Repetition Time

Image Size (Width and Height)

Slice Thickness

Reconstruction Diameter (Field of View)

Receiver Coil

From this information determine the in-plane resolution and the volume of an individual voxel.

Exercise 2a

Using Program\_02a perform mono- and bi-exponential fitting on the diffusion weighted data provided within the program using 4 data points and then 20 data points. Using your results determine which fitting routine is most appropriate for each set of data and provide a rationale for your answer.

*It appears from the data that has been acquired from the MATLAB script the Bi-exponential fitting is more accurate than Mono- exponential fitting in this instance.  
Bi-exponential fitting on average has less average distance between itself and the points of data than Mono-exponential fitting (0.0011 to 0.0013 and 0.0010 to 0.0014 respectively). Bi-exponential fitting also on average has lower error than Mono-exponential fitting by magnitudes (0.0 to 6373.0 to 132.3 to 3469.2 respectively), in fact when using 4 points Bi-exponential incurs no error at all.  
  
In the case of Bi-exponential fitting it appears that it operates better with fewer data points, this is because the error incurred when using more points increased exponentially, however it does also appear that the distance between the line and points increased marginally.  
In the case of Mono-exponential fitting it appears that it operates better with many data points, this is because the error halved when using more data points, also the distance between the line and the data points decreased slightly.  
  
Thus, it would be safe to assume that when using few data points to plot a curve it is best to use Bi-exponential fitting and when using many data points to plot a curve it would be best to use Mono-exponential fitting.*

Exercise 2b

Using Program\_02b determine the most appropriate polynomial order to fit the data provided within the program. You may need to run the program a number of times to evaluate this.

*In this case by observing the data which the polynomial curve is to be plotted over it can quickly be determined that the optimum polynomial order to use would be a polynomial curve of order 4, this is because from looking at the data it can be seen that the data has 3 inflection points, a polynomial curve of order 4 also has 3 inflection points. The number of inflection points that a polynomial curve will have is determined by taking the order of the polynomial (in this case 4) and subtracting 1 from it, a 1st order polynomial will be a straight line with no inflection points, a 2rd order polynomial will have 1 inflection point and a 99th order polynomial will have 98 inflection points.  
  
However, it can also be determined by looking at the data that a 4th order polynomial is ideal. Straight away the 1st order polynomial can be ignored, this is because as discussed previously it is a straight line and as such is not fit for this purpose, as can be seen its R squared is very low and its AIC is very high.*

*The 5h order polynomial and any orders above can also be ignored, this can be seen in the data as, every polynomial between the 2nd order and the 4th order has an AIC which decreases and then on the 5th order the AIC starts to increase again.  
When using AIC polynomials it is best to not use an order which is too high as the algorithm will start to force the curve to fit the data points while still maintaining inflection points, this can lead to overfitting the polynomial to the data and increasing error.*

*From the polynomials remaining the 4th order polynomial is the polynomial which has the highest R squared.*

Exercise 3

Using Program\_03 investigate the shape parameters of the lesion found in the file *MRI\_BRAIN*

The brightness and contrast can be altered using the slider bars displayed beneath the image, and the image can be thresholded using the slider bar on the right. Once you are satisfied with the threshold value use the push button ‘Select Tumour Pixel’ to select a pixel within the tumour so that all other thresholded areas are removed. Parameters can then be calculated using the ‘Calculate Parameters’ push button. Investigate how the shape parameters vary by exploring between threshold values of 3000 and 4000. Which parameter has the lowest percentage variation when changing the threshold level?

*When starting at threshold level 3000 the segmentation algorithm still includes a lot of surrounding pixels in the region of interest with the lesion which we are trying to isolate. This can be seen by comparing the parameters of the segmentation using a threshold level of 3000 and the segmentation using a threshold level of 3091, the size of the 3000 segmentation is dramatically larger than the 3091 segmentation with 153 compared to 145, however when comparing the 3091 segmentation to the 3182 segmentation it can be seen that no change in size occurs with 145 to 145. The other parameters are similar.  
  
Conversely, it can be seen that when starting at the largest threshold value of 3910 and comparing to smaller threshold values that at this point the algorithm is starting to remove parts of the lesion from the segment, this can be seen by comparing the Elongatedness of all segmentations, it can be seen that from segmentation 3091 to segmentation 3728 that all segmentations share the same value of 1.000 meaning that the proportion of the segmentation does not change dramatically, thus it can be assumed that the best threshold value for this segmentation is within this range.  
  
On a tangent, the lowest percentage change of any parameter is the Elongatedness which as mentioned previously is unchanging between 3091 and 3728 meaning that there is a 0% change in this range, between all ranges there is a 12.5% change from highest value to lowest value.  
  
If we then look at the other parameters to try and determine the best threshold value, if we take the highest Circularity and Convexity to mean the best threshold value that value in this instance would be 3273.*

Exercise 4

With Program\_04 perform k-means clustering on the images *MRI\_PRE* and *MRI\_POST*. Segment the data into grey matter, white matter and enhancing lesion. Deduce the tumour volume.

*By segmenting the data into two clusters the proportion of grey matter to white matter and enhancing lesion can be found 7010.0 mm^3 to 7338.0 mm^3, if then the data is then segmented into 3 clusters the volume of the enhancing lesion can be found 152.0 mm^3, the other clusters appear to include extraneous data an so cannot be used to deduce the volume of the grey matter to white matter. Then from this data we can estimate the volume of the white matter by taking the volume of the enhancing lesion and subtracting it from the volume of the white matter and enhancing lesion thus deducing that the volume of grey matter is 7010.0 mm^3, the volume of white matter is 7186.0 mm^3 and the volume of the enhancing lesion is 152.0 mm^3.*

Exercise 5

Using Program\_05 calculate the 2nd order texture features angular second moment and entropy, for both *CASE A* and *CASE B*, at a range of grey levels from 4 to 256.

According to your results which tumour is more heterogeneous?

Illustrate and discuss the variation in the texture parameters with increasing number of grey levels.

*Heterogeneous is the antonym of homogenous, homogenous means unchanging and heterogeneous means rapidly changing or changing by a great magnitude. Thus, a heterogeneous tumour would be a tumour where in this case its grey levels are changing rapidly or changing by a great magnitude.  
A tumour which was malignant would be heterogeneous and be extension would be less circular etc.  
  
In this case, case B or the second tumour is less heterogeneous than the first tumour case A, thus, tumour A is worse than tumour B.  
  
The heterogeneity of a tumour can be inferred from in this case the change in the ASM and Entropy over the many numbers of grey levels and can also be estimated by comparing the level of each tumours ASM and Entropy to each other for each given grey level.  
  
In this case over nearly every grey level the magnitude of case A’s ASM and Entropy is greater than the magnitude of case B’s ASM and Entropy, for 4 grey levels 0.1270 to 0.1235 and 3.2663 to 3.2840 and fro 256 grey levels 0.0011 to 0.0030 and 9.8407 to 8.4088 respectively.  
The percentage change of case A’s ASM from 4 to 256 grey levels is an 11545% decrease and of case A’s ASM from 4 to 256 grey levels is a 307% increase. Whereas for case B it has a 4117% decrease and a 256% increase.  
  
As can be seen an 11545% decrease is greater than a 4117& decrease and a 307% increase is greater than a 256% increase, thus case A is changing more than case B and as such is more heterogeneous.*