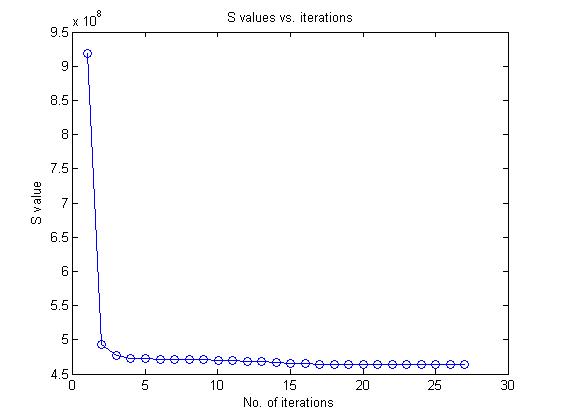
Convergence criteria

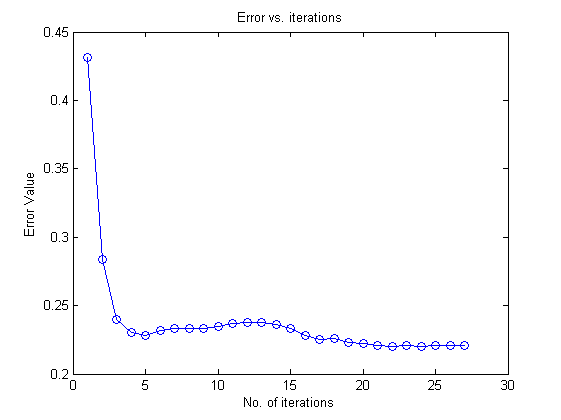
Here xi denotes the example point in space and denotes the cluster centre to which xi is assigned.

The convergence criteria used to epsilon bound on the sum of the squares of the distance from the mean of the assigned cluster. The epsilon was chosen to be around 0.001. The algorithm converged in approximately 25-27 iterations depending on the initial random cluster assignment.



The initial S value is after the random assignment of cluster and the value decreases successively.

After one iteration of the algorithm, the value decreases substantially and then slowly converges according to the convergence criteria at about 25-27 iterations. Since the graph on the next page shows that the error reduces, therefore the convergence criteria performs well.



The error in clustering mostly decreases with increase in number of iterations. It increases in between which may be due to misclassification of some boundary point. Some outliers may have led to a cluster becoming skewed and misclassification of certain data points. But in successive iterations, due to correct classification of outliers, the boundary points might have been correctly classified. Ultimately, the error reduces with increase in number of iterations. Accuracy in range of 76-78% has been observed.