EM Algorithm

Geoffrey Ulman
Homework 8
CSI873

November 2011

1 Results

Running the EM algorithm on the provided 2000 values (known to be drawn from two Gaussian distributions with variance 1.0) resulted in mean estimates of 0.45 and -0.74. The evolution of the two mean estimates as the EM algorithm progressed is plotted in Figure 1. A histogram of the full data set is provided in Figure 2 as a sanity check.

Java (version 1.6.0_27) was used to implement the EM algorithm. The code is available as a Subversion repository on Google Code at `http://code.google.com/p/csi873/`. Compiling and running the code requires the Java build tool Maven (`http://maven.apache.org/`).

References

[1] Tom M. Mitchell, *Machine Learning*, WCB McGraw-Hill, Boston, 1997.

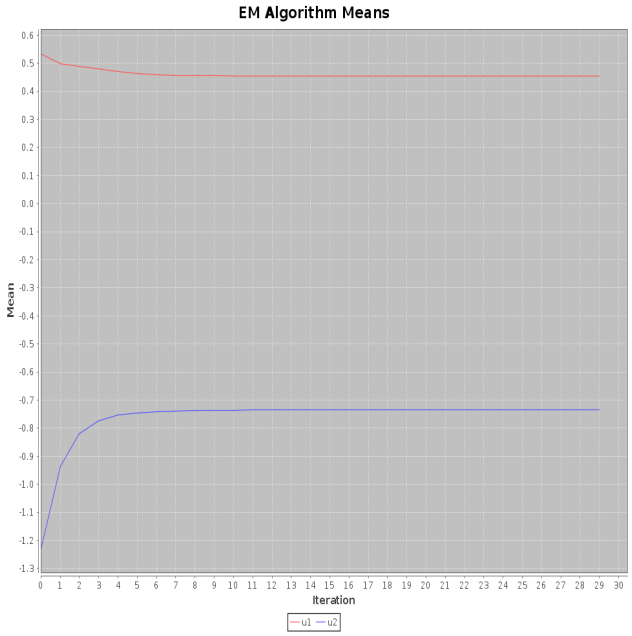

Figure 1: Evolution of Mean Estimates using EM Algorithm

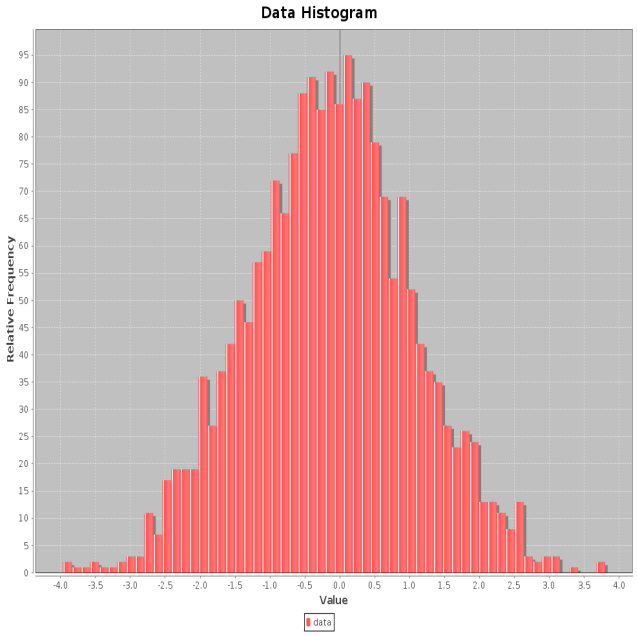

Figure 2: EM Data Set Geoffrey.txt