Subject: Re: clustering
Posted by little davey on Mon, 23 Jul 2007 17:14:49 GMT
View Forum Message <> Reply to Message

Is it the case that you MUST use standardize() before you call
CLUST_WTS()? The documentation does not say so, but I suspect from
the code, and, I actually tried it with the initial poster's data, and
got 4 clusters (he wanted 5, but this would appear to be a tough data
set to cluster, as variables "2" and "3" are close to each other).

As I posted an hour ago, part of the source code for CLUST_WTS() is:

```
  ;Normalized uniformly random cluster weights.
  ;;Weights = RANDOMU(SEED, N_Variables, N_Clusters) + Zero
      av1 = average(array[0,*])
      av2 = average(array[1,*])
  Weights = RANDOMU(SEED, N_Variables, N_Clusters) + Zero
  for k = 0L, N_Clusters-1 do Weights[*,k] = $
    (Weights[*,k] / TOTAL(Weights[*,k])) * Variable_Wts
```

However, the variables AV1 and AV2 are NOT USED ANYWHERE IN THE CODE,
so I suspect that the data is not "normalized" correctly in
CLUST_WTS. The use of STANDARDIZE() may be necessary for CLUST_WTS to
work.

-- Dave K --