
Subject: Re: clustering

Posted by [nivedita.raghunath](#) on Thu, 19 Jul 2007 16:12:44 GMT

[View Forum Message](#) <> [Reply to Message](#)

Here is a subset of my data.

IDL> help,pos1

POS1 FLOAT = Array[7, 53]

IDL> print,pos1

0.910300	0.413400	-0.0221000	0.00300000	-150.250
129.510	-13.0400			
0.910200	0.413400	-0.0223000	0.00370000	-150.280
129.460	-13.0200			
0.910200	0.413500	-0.0228000	0.00360000	-150.300
129.400	-13.1300			
0.910200	0.413400	-0.0231000	0.00310000	-150.190
129.520	-13.0700			
0.910200	0.413400	-0.0226000	0.00320000	-150.220
129.580	-13.0800			
0.910200	0.413600	-0.0224000	0.00460000	-150.510
129.040	-13.1000			
0.910200	0.413500	-0.0221000	0.00250000	-150.210
129.560	-13.0000			
0.910200	0.413500	-0.0223000	0.00340000	-150.310
129.420	-13.1000			
0.910200	0.413500	-0.0225000	0.00350000	-150.160
129.620	-13.0900			
0.910200	0.413500	-0.0224000	0.00240000	-150.090
129.720	-13.0100			
0.930600	0.365500	-0.0216000	0.00170000	-147.800
125.760	-16.7500			
0.930500	0.365600	-0.0220000	0.000900000	-147.650
125.160	-16.6800			
0.930500	0.365700	-0.0222000	0.00230000	-147.930
125.370	-16.8100			
0.930500	0.365700	-0.0217000	0.00280000	-148.090
125.750	-16.8600			
0.930400	0.365800	-0.0225000	0.00240000	-147.800
125.400	-16.8200			
0.930400	0.365800	-0.0213000	0.00430000	-148.490
124.950	-16.7800			
0.930400	0.365800	-0.0220000	0.00210000	-147.910
126.000	-16.7200			
0.930400	0.365800	-0.0220000	0.00200000	-147.830
125.560	-16.6900			
0.930400	0.365900	-0.0216000	0.00250000	-148.080
125.490	-16.7700			

0.930400	0.365800	-0.0224000	0.00230000	-147.870
125.980	-16.6200			
0.897600	0.439600	-0.0331000	0.00790000	-147.060
130.970	-6.02000			
0.897600	0.439500	-0.0334000	0.00720000	-146.790
130.520	-6.13000			
0.897500	0.439600	-0.0337000	0.00770000	-146.820
130.660	-6.13000			
0.897500	0.439600	-0.0328000	0.00750000	-147.160
130.790	-6.13000			
0.897600	0.439600	-0.0331000	0.00680000	-146.860
130.570	-6.07000			
0.897600	0.439600	-0.0335000	0.00700000	-146.830
130.660	-6.12000			
0.897600	0.439500	-0.0326000	0.00750000	-147.090
130.870	-6.08000			
0.897600	0.439600	-0.0327000	0.00750000	-146.880
130.610	-6.14000			
0.897600	0.439500	-0.0336000	0.00810000	-146.980
130.560	-6.25000			
0.897600	0.439500	-0.0331000	0.00800000	-147.130
130.820	-6.19000			
0.897500	0.439600	-0.0332000	0.00800000	-147.000
130.600	-6.25000			
0.871700	0.488800	-0.0332000	0.0102000	-146.260
133.480	-1.14000			
0.871600	0.488900	-0.0330000	0.0111000	-146.390
133.540	-1.29000			
0.871600	0.488900	-0.0347000	0.00920000	-145.690
132.630	-1.26000			
0.871700	0.488800	-0.0337000	0.0103000	-146.100
133.330	-1.44000			
0.871700	0.488700	-0.0336000	0.0104000	-146.310
133.610	-1.58000			
0.871700	0.488800	-0.0340000	0.00950000	-145.820
132.840	-1.33000			
0.872000	0.488200	-0.0335000	0.00960000	-146.040
133.140	-1.95000			
0.872000	0.488200	-0.0330000	0.00820000	-145.910
133.210	-1.83000			
0.872000	0.488300	-0.0333000	0.0100000	-146.040
133.110	-1.82000			
0.872100	0.488200	-0.0330000	0.00880000	-146.000
133.150	-1.83000			
0.872000	0.488200	-0.0335000	0.00900000	-145.910
133.210	-1.85000			
0.873000	0.487300	-0.0227000	0.000700000	-143.720
132.260	-6.08000			

```

    0.872900  0.487300 -0.0230000 0.000100000 -143.630
132.350 -6.07000
    0.872900  0.487300 -0.0235000 0.000500000 -143.560
132.370 -6.14000
    0.872900  0.487300 -0.0234000 -0.000300000 -143.430
132.520 -6.15000
    0.872900  0.487300 -0.0231000 0.000700000 -143.670
132.280 -6.15000
    0.872900  0.487300 -0.0237000 0.000200000 -143.480
132.430 -6.07000
    0.872900  0.487300 -0.0231000 0.000500000 -143.550
132.450 -6.03000
    0.872900  0.487300 -0.0241000 -0.000200000 -143.440
132.450 -6.11000
    0.872900  0.487300 -0.0237000 0.000400000 -143.470
132.490 -6.05000
    0.873000  0.487300 -0.0228000 0.000600000 -143.700
132.270 -6.03000
    0.872900  0.487300 -0.0235000 -0.000200000 -143.430
132.450 -6.10000

```

```

IDL> weights=clust_wts(pos1,n_clusters=5)
IDL> print,weights
    0.159265  0.119451  0.113155  0.180601  0.0680267
0.243488  0.116014
    0.874568  0.483835 -0.0256644  0.00231699  -144.219
132.388 -5.40240
    0.113501  0.127323  0.0985566  0.247231  0.225678
0.0779656  0.109745
    0.238006  0.236222  0.127174  0.0261984  0.266028
0.0180832  0.0882878
    0.0301962  0.232814  0.209770  0.146116  0.235975
0.134589  0.0105386

```

```

IDL> result=cluster(pos1,weights,n_clusters=5)
IDL> print,result(uniq(result))
1

```

The 5 clusters are pretty distinct but "cluster" does a hopeless job identifying them. I tried scaling the data but that too finds 3 clusters in the end. Any ideas?

On Jul 18, 3:07 pm, Conor <cmanc...@gmail.com> wrote:
> On Jul 17, 11:36 am, nivedita.raghun...@gmail.com wrote:
>

>> Hi all,
>
>> I am trying to cluster a (7 x n) array with n_clusters=5. Visually I
>> can see 5 distinct clusters, but when I do clust_wts the cluster
>> centroids don't end up right. No matter what options I give, clust_wts
>> refuses to find the clusters.
>
>> Any idea on whats going on ?
>
>> Thanks,
>> Nivedita
>
> No idea. More information might be helpful. It's quite possible
> though that the the clust_wts algorithm just doesn't work for your
> particular data set, at least not as well as you apparently want it to.
