Subject: Re: Need Some Advice on Seperating Out Some Data
Posted by rdellsy on Wed, 09 Aug 2006 20:13:12 GMT
View Forum Message <> Reply to Message

Thanks for that. I took it, and played around with it a bit to get it
to work. [Errors I found were: x and y don't concatinate in the line
'array=transpose([[x],[y]])' and I found I had to comment away the
/ISOTROPIC in the plotting.) Unfortunately, it seems that cluster
seperates on a purely 1 dimensional basis. I tried discarding the
histogram related code in favor of a much simpler system in case that
was the problem, and it still didn't work. If you look at the data set
I provided, the problem should be self evident.

Incidentally, I replaced everything from
h=histogram(c,reverse_indices=ri) down to the second to last line with:
--
plot,x,y,psym=2
bmax=max(array[0,*],maxsubsc)
goodc=c[maxsubsc]
keep=where(c[*] eq goodc)
--
I feel that my code may be a tad more efficient, though I don't know
how efficient the WHERE command is.

Anywho, I'm looking CLUSTER_TREE right now, which shows some more
promise. If I understand it correctly, it works using distance appart,
not coordinates which is a bit more useful, I think, for my problem.
I'm just not sure how I can take the output and turn it into a set of
clusters.

Thanks for all the help!
- Rob

JD Smith wrote:
> On Tue, 08 Aug 2006 15:20:13 -0700, rdellsy wrote:
>
>> I'm a tad confused about what you're suggesting. I'll try and work it
>> out, but I'm still fairly new to IDL, so if you could give an IDL or
>> pseudo-code example of what you're trying to explain, I would
>> appreciate. If that's too much work, I understand, and I'll just try to
>> puzzle it out on my own.
>
> You might find much of what you need in the HISTOGRAM tutorial:
>
> http://www.dfanning.com/tips/histogram_tutorial.html
>
> But before you go that route, you might first try the CLUSTER function
> in IDL (which I just read up on). Here's an example using a fake

```
>  clustered data set with 5 clusters.  You'll probably have to experiment
>  with the number of clusters.
>
>  JD
>
>   tvlct,[255,0,0,0,255,255],[0,255,0,255,255,0],[0,0,255,255,0 ,255],1
>  n_clust=5
>
>  ;; Make some flake clustered data
>  if n_elements(x) ne 0 then begin
>     n=1000
>     clust_fwhm=.2
>     cposx=randomu(sd,n_clust) & cposy=randomu(sd,n_clust)
>     cind=fix(randomu(sd,n)*n_clust)
>
>     x=clust_fwhm
>     fac=2*sqrt(2*alog(2))
>     x=randomn(sd,n)*clust_fwhm/fac+cposx[cind]
>     y=randomn(sd,n)*clust_fwhm/fac+cposy[cind]
>  endif
>
>  array=transpose([[x],[y]])
>  w=clust_wts(array,N_CLUSTERS=n_clust)
>  c=cluster(array,w)
>  h=histogram(c,REVERSE_INDICES=ri)
>  nh=n_elements(h)
>
>  plot,x,y,PSYM=4,/ISOTROPIC
>
>  cen=make_array(2,nh,VALUE=!VALUES.F_NAN)
>  for i=0,nh-1 do begin
>     if ri[i+1] eq ri[i] then continue
>     take=ri[ri[i]:ri[i+1]-1]
>     oplot,x[take],y[take],PSYM=4,COLOR=i+1
>     cen[0,i]=[mean(x[take]),mean(y[take])]
>  endfor
>
>  ;; Find the lower right cluster
>  void=max(cen[0,*]-cen[1,*],lrc,/NAN)
>
>  ;; Highlight it
>  keep=ri[ri[lrc]:ri[lrc+1]-1]
>  oplot,x[keep],y[keep],PSYM=6,SYMSIZE=2
>
>  END
```