Subject: Clustering x,y coordinates with IDL?
Posted by Maxwell Peck on Mon, 29 Mar 2010 09:44:02 GMT
View Forum Message <> Reply to Message

Hi All, This is not strictly an IDL question but I'm hoping it's something simple to do and I'm not missing something painfully obvious!

I have a set of x,y values e.g.

x = [34.5,36.7,35.6,80.5,81.2,79.3]
y = [50.6,51.2,53.2,48.5,50.3,51.2]

The values will be sparsely clustered overall but tightly within an approximate 10x10 box. What I'd like to do is replace each clusters values with the average x,y value. This doesn't have to be 'perfect', the edges aren't that important.

My initial thought was to use hist_2d or hist_nd to try and calculate the 2d histogram with a binsize of 10 and use this to average the values but I can't get it to work. My other thought was using griddata or something similar, or at worst generating an approximate binary image and running a window over it.

Any advice or suggestions would be appreciated.

Regards

Max

Subject: Re: Clustering x,y coordinates with IDL?
Posted by Mort Canty on Mon, 29 Mar 2010 15:38:08 GMT
View Forum Message <> Reply to Message

Am 29.03.2010 11:44, schrieb Maxwell Peck:
> Hi All, This is not strictly an IDL question but I'm hoping it's
> something simple to do and I'm not missing something painfully
> obvious!
>
> I have a set of x,y values e.g.
>
> x = [34.5,36.7,35.6,80.5,81.2,79.3]
> y = [50.6,51.2,53.2,48.5,50.3,51.2]
>
> The values will be sparsely clustered overall but tightly within an
> approximate 10x10 box. What I'd like to do is replace each clusters
> values with the average x,y value. This doesn't have to be 'perfect',

> the edges aren't that important.
>
> My initial thought was to use hist_2d or hist_nd to try and calculate
> the 2d histogram with a binsize of 10 and use this to average the
> values but I can't get it to work. My other thought was using griddata
> or something similar, or at worst generating an approximate binary
> image and running a window over it.
>
> Any advice or suggestions would be appreciated.
>
> Regards
>
> Max
>

The IDL function CLUSTER() implements k-means clustering. I think that's what you want.

Mort

Subject: Re: Clustering x,y coordinates with IDL?
Posted by Maxwell Peck on Mon, 29 Mar 2010 20:20:17 GMT
View Forum Message <> Reply to Message

On Mar 30, 2:38 am, Mort Canty <m.ca...@fz-juelich.de> wrote:
> Am 29.03.2010 11:44, schrieb Maxwell Peck:
>
>
>
>> Hi All, This is not strictly an IDL question but I'm hoping it's
>> something simple to do and I'm not missing something painfully
>> obvious!
>
>> I have a set of x,y values e.g.
>
>> x = [34.5,36.7,35.6,80.5,81.2,79.3]
>> y = [50.6,51.2,53.2,48.5,50.3,51.2]
>
>> The values will be sparsely clustered overall but tightly within an
>> approximate 10x10 box. What I'd like to do is replace each clusters
>> values with the average x,y value. This doesn't have to be 'perfect',
>> the edges aren't that important.
>
>> My initial thought was to use hist_2d or hist_nd to try and calculate
>> the 2d histogram with a binsize of 10 and use this to average the
>> values but I can't get it to work. My other thought was using griddata

>> or something similar, or at worst generating an approximate binary
>> image and running a window over it.
>
>> Any advice or suggestions would be appreciated.
>
>> Regards
>
>> Max
>
> The IDL function CLUSTER() implements k-means clustering. I think that's
> what you want.
>
> Mort

Thanks Mort. I had seen this I am concerned though about how it will deal with single/a few points as opposed to real 'clusters'. Also the distance with which points are considered to be a cluster as well is concerning. I will give it a bash anyway and see how it looks. I had thought there must be a nice way to do it with histogram but I just can't get it to work.

Regards
Nax
