Subject: Re: Pair Counts in an Annulus, for large data sets
Posted by enod on Sat, 11 Nov 2006 11:51:25 GMT
View Forum Message <> Reply to Message

Maybe you can try WHERE((seperation gt (r)) and (seperation lt (r +
deltar_r)).

Tian

On Nov 11, 1:52 pm, fatcat3...@gmail.com wrote:
> Hi There
>
> I have a large data set (~350,000 galaxies) of x,y points. For a given
> radius R [R=sqrt(x^2+y^2)], I need to count the total number of pairs
> in the annulus R+deltaR. That is, I choose a given data point as my
> center, then count the number of points that lie inside that annulus. I
> then do this for each of my data points to get the total number of
> pairs. A simplified version of the code I'm using now is as follows:
>
> ****************
> n = n_elements(x) ; number of data points
> seperation = fltarr(n) ; the seperation between data points
>
> for i=0L,n-1 do begin
> seperation = sqrt((x - x[i])^2 + (y - y[i])^2) ;distance between the
> two points, centering on point "i"
> seperation[i] = 999 ;simply because I don't want it to count itself
> as a pair
>
> if_inside = ((seperation gt (r)) and (seperation lt (r + deltar_r))
> ;has value "1" for points which lie inside, "0" for those outside
> counter = counter + total(if_sep) ;count up the number of pairs
> endfor
>
> num_pairs = counter / 2 ; since I don't want to count everything twice
> ******************
>
> I've tried my best to avoid the urge to put lots of for loops
> everywhere (you should have seen it before!), but I just don't know how
> to make it drastically more efficient. There must be a way though,
> because the computations for my code are just ridiculous.... Is there a
> way to eliminate that nasty loop I have, which would help things?
>
> Any help you can give would be greatly appreciated. I'm very new to
> IDL, as you surely know. I'm and undergrad, too.
>

> Tara.

Subject: Re: Pair Counts in an Annulus, for large data sets
Posted by David Fanning on Sat, 11 Nov 2006 12:28:22 GMT
View Forum Message <> Reply to Message

Tara writes:

> I have a large data set (~350,000 galaxies) of x,y points. For a given
> radius R [R=sqrt(x^2+y^2)], I need to count the total number of pairs
> in the annulus R+deltaR. That is, I choose a given data point as my
> center, then count the number of points that lie inside that annulus. I
> then do this for each of my data points to get the total number of
> pairs. A simplified version of the code I'm using now is as follows:
>
> I've tried my best to avoid the urge to put lots of for loops
> everywhere (you should have seen it before!), but I just don't know how
> to make it drastically more efficient. There must be a way though,
> because the computations for my code are just ridiculous.... Is there a
> way to eliminate that nasty loop I have, which would help things?
>
> Any help you can give would be greatly appreciated. I'm very new to
> IDL, as you surely know. I'm and undergrad, too.

Well, then, you are about to learn something about the most
mysterious IDL command of all: HISTOGRAM. For that is
exactly what you are trying to do. HISTOGRAM counts
how many things are in each "bin" of a particular size.
And it can even tell you which of the things you are
counting are stored in each bin.

I think your problem can be solved something like this

```
r = SQRT(x^2+y^2)
h = HISTOGRAM(r, BINSIZE=deltaR, REVERSE_INDICES=ri)
```

The famous Histogram Tutorial will provide all the information
you need to get beyond this:

http://www.dfanning.com/tips/histogram_tutorial.html

Cheers,

David
--
David Fanning, Ph.D.
Fanning Software Consulting, Inc.

Subject: Re: Pair Counts in an Annulus, for large data sets
Posted by fatcat3131 on Thu, 23 Nov 2006 04:58:20 GMT
View Forum Message <> Reply to Message

hi there, me again...

i've been tinkering with things, and i simply don't understand the histogram principle (i even read the tutorial!). would you mind explaining in slightly more detail?

it's really not clicking with me how setting the binsize to Delta-R would allow me to count the points which lie in the distances (R to Delta-R) from a point P, and do this for all N points. i am interested in just one Delta-R, and counting the point-point pairs for this Delta-R.

missing something,

Tara

David Fanning wrote:
> Tara writes:
>
>> I have a large data set (~350,000 galaxies) of x,y points. For a given
>> radius R [R=sqrt(x^2+y^2)], I need to count the total number of pairs
>> in the annulus R+deltaR. That is, I choose a given data point as my
>> center, then count the number of points that lie inside that annulus. I
>> then do this for each of my data points to get the total number of
>> pairs. A simplified version of the code I'm using now is as follows:
>>
>> I've tried my best to avoid the urge to put lots of for loops
>> everywhere (you should have seen it before!), but I just don't know how
>> to make it drastically more efficient. There must be a way though,
>> because the computations for my code are just ridiculous.... Is there a
>> way to eliminate that nasty loop I have, which would help things?
>>
>> Any help you can give would be greatly appreciated. I'm very new to
>> IDL, as you surely know. I'm and undergrad, too.
>
> Well, then, you are about to learn something about the most
> mysterious IDL command of all: HISTOGRAM. For that is
> *exactly* what you are trying to do. HISTOGRAM counts
> how many things are in each "bin" of a particular size.

> And it can even tell you which of the things you are
> counting are stored in each bin.
>
> I think your problem can be solved something like this
>
>   r = SQRT(x^2+y^2)
>   h = HISTOGRAM(r, BINSIZE=deltaR, REVERSE_INDICES=ri)
>
> The famous Histogram Tutorial will provide all the information
> you need to get beyond this:
>
>   http://www.dfanning.com/tips/histogram_tutorial.html
>
> Cheers,
>
> David
> --
> David Fanning, Ph.D.
> Fanning Software Consulting, Inc.
> Coyote's Guide to IDL Programming: http://www.dfanning.com/
> Sepore ma de ni thui. ("Perhaps thou speakest truth.")
