Subject: Re: matching coordinates
Posted by David Fanning on Sun, 29 Jul 2007 13:53:08 GMT
View Forum Message <> Reply to Message

jradavenport@gmail.com writes:

> Hello, I'm a long time lurker of this group, and finally have a
> problem I can't solve efficiently enough! I have coordinates for stars
> (RA and DEC for us astronomers out there), from two sources. i.e.
> (x1,y1) and (x2,y2). These lists are huge, like 150k for one and
> 5million in the other (a lot of stars!). I need to find the matches
> from these catalogs within a tolerance. I've tried a few programs
> I've found online (close_match_radec for instance) and they have not
> given me results I believe (multiple matches and such). Here is some
> horrible code I wrote just now which solves the problem VERY slowly:

Here is an article that might help:

http://www.dfanning.com/code_tips/matchlists.html

For background, you might want to read the precursor
article, and one of my personal favorites:

http://www.dfanning.com/code_tips/slowloops.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.")

Subject: Re: matching coordinates
Posted by Anthony[1] on Mon, 30 Jul 2007 07:41:44 GMT
View Forum Message <> Reply to Message

On Jul 29, 10:01 am, jradavenp...@gmail.com wrote:
> I need to find the matches
> from these catalogs within a tolerance.

I tend to use TOPCAT for this kind of thing:

http://www.star.bris.ac.uk/~mbt/topcat/

But if there is a way to do the same thing equally well in IDL I'd be

interested to know.

Regards,

Anthony

Subject: Re: matching coordinates
Posted by Anthony[1] on Mon, 30 Jul 2007 07:41:59 GMT
View Forum Message <> Reply to Message

On Jul 29, 10:01 am, jradavenp...@gmail.com wrote:
> I need to find the matches
> from these catalogs within a tolerance.

I tend to use TOPCAT for this kind of thing:

http://www.star.bris.ac.uk/~mbt/topcat/

But if there is a way to do the same thing equally well in IDL I'd be
interested to know.

Regards,

Anthony

Subject: Re: matching coordinates
Posted by jradavenport on Mon, 30 Jul 2007 10:36:30 GMT
View Forum Message <> Reply to Message

On Jul 29, 7:53 am, David Fanning <n...@dfanning.com> wrote:
> jradavenp...@gmail.com writes:
>> Hello, I'm a long time lurker of this group, and finally have a
>> problem I can't solve efficiently enough! I have coordinates for stars
>> (RA and DEC for us astronomers out there), from two sources. i.e.
>> (x1,y1) and (x2,y2). These lists are huge, like 150k for one and
>> 5million in the other (a lot of stars!). I need to find the matches
>> from these catalogs within a tolerance. I've tried a few programs
>> I've found online (close_match_radec for instance) and they have not
>> given me results I believe (multiple matches and such). Here is some
>> horrible code I wrote just now which solves the problem VERY slowly:
>
> Here is an article that might help:
>
> http://www.dfanning.com/code_tips/matchlists.html
>

> For background, you might want to read the precursor
> article, and one of my personal favorites:
>
> http://www.dfanning.com/code_tips/slowloops.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.")

Thanks David, Big fan of your site and of yours and JD Smith's code...
I guess it really is time to learn that Histogram function (something
I've put off for over a year!)

I'm turning that example code into a workable program, but am confused
on where the results are. Which Variable do I want to return with the
answers?

Subject: Re: matching coordinates
Posted by jradavenport on Mon, 30 Jul 2007 11:12:23 GMT
View Forum Message <> Reply to Message

On Jul 30, 4:36 am, jradavenp...@gmail.com wrote:
> On Jul 29, 7:53 am, David Fanning <n...@dfanning.com> wrote:
>
>
>
>> jradavenp...@gmail.com writes:
>>> Hello, I'm a long time lurker of this group, and finally have a
>>> problem I can't solve efficiently enough! I have coordinates for stars
>>> (RA and DEC for us astronomers out there), from two sources. i.e.
>>> (x1,y1) and (x2,y2). These lists are huge, like 150k for one and
>>> 5million in the other (a lot of stars!). I need to find the matches
>>> from these catalogs within a tolerance. I've tried a few programs
>>> I've found online (close_match_radec for instance) and they have not
>>> given me results I believe (multiple matches and such). Here is some
>>> horrible code I wrote just now which solves the problem VERY slowly:
>
>> Here is an article that might help:
>
>> http://www.dfanning.com/code_tips/matchlists.html
>
>> For background, you might want to read the precursor

>> article, and one of my personal favorites:
>
>> http://www.dfanning.com/code_tips/slowloops.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.")
>
> Thanks David, Big fan of your site and of yours and JD Smith's code...
> I guess it really is time to learn that Histogram function (something
> I've put off for over a year!)
>
> I'm turning that example code into a workable program, but am confused
> on where the results are. Which Variable do I want to return with the
> answers?

Actually, scratch that. I believe it's the min_pos and min_dist
variables. The only other problem I'm having is it's telling me in
line 40 (including a program begin line)

% HISTOGRAM: Illegal binsize or max/min.

This is for the line:

```
;; Dual HISTOGRAM method, loop by repeat count in bins
h2 = histogram(h[b],MIN=1,REVERSE_INDICES=ri2)
```

Anybody have ideas who knows about Histogram?

Subject: Re: matching coordinates
Posted by Conor on Mon, 30 Jul 2007 14:58:20 GMT

View Forum Message <> Reply to Message

On Jul 30, 7:12 am, jradavenp...@gmail.com wrote:
> On Jul 30, 4:36 am, jradavenp...@gmail.com wrote:
>
>
>
>> On Jul 29, 7:53 am, David Fanning <n...@dfanning.com> wrote:
>
>>> jradavenp...@gmail.com writes:
>>>> Hello, I'm a long time lurker of this group, and finally have a
>>>> problem I can't solve efficiently enough! I have coordinates for stars

>>>> (RA and DEC for us astronomers out there), from two sources. i.e.
>>>> (x1,y1) and (x2,y2). These lists are huge, like 150k for one and
>>>> 5million in the other (a lot of stars!). I need to find the matches
>>>> from these catalogs within a tolerance. I've tried a few programs
>>>> I've found online (close_match_radec for instance) and they have not
>>>> given me results I believe (multiple matches and such). Here is some
>>>> horrible code I wrote just now which solves the problem VERY slowly:
>
>>> Here is an article that might help:
>
>>>   http://www.dfanning.com/code_tips/matchlists.html
>
>>> For background, you might want to read the precursor
>>> article, and one of my personal favorites:
>
>>>   http://www.dfanning.com/code_tips/slowloops.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.")
>
>> Thanks David, Big fan of your site and of yours and JD Smith's code...
>> I guess it really is time to learn that Histogram function (something
>> I've put off for over a year!)
>
>> I'm turning that example code into a workable program, but am confused
>> on where the results are. Which Variable do I want to return with the
>> answers?
>
> Actually, scratch that. I believe it's the min_pos and min_dist
> variables. The only other problem I'm having is it's telling me in
> line 40 (including a program begin line)
>
> % HISTOGRAM: Illegal binsize or max/min.
>
> This is for the line:
>   ;; Dual HISTOGRAM method, loop by repeat count in bins
>   h2 = histogram(h[b],MIN=1,REVERSE_INDICES=ri2)
>
> Anybody have ideas who knows about Histogram?

As an alternate solution, here's what I use:

astro.ufl.edu/~cmancone/pros/qfind.pro

I wrote this code as a result of the discussion here:

http://groups.google.com/group/comp.lang.idl-pvwave/browse_t
hread/thread/629cbb2a852c5371/6ada6d1659bc55a7?hl=en#6ada6d1 659bc55a7

You would call:

result = qfind(x1,y1,x2,y2,posshift=tolerance)

'result' would be a 2xn array where n is the total number of matched
stars. result is essentially the equivelent of two where functions.
So for instance:

abs(x1[result[0,*]] - x2[result[1,*]]) is always less than tolerance

x1,y1,x2,y2 all need to be row vectors. If you pass them as column
vectors, they will automatically be transposed (which, now that I
think about it, might be bad if the rest of your program expects
column vectors). I'm not sure how it compares to the other for speed
(probably worse), but I'd be curious.

Subject: Re: matching coordinates
Posted by JD Smith on Tue, 31 Jul 2007 00:20:55 GMT
View Forum Message <> Reply to Message

On Mon, 30 Jul 2007 11:12:23 +0000, jradavenport wrote:

>
> % HISTOGRAM: Illegal binsize or max/min.
>
> This is for the line:
> ;; Dual HISTOGRAM method, loop by repeat count in bins
> h2 = histogram(h[b],MIN=1,REVERSE_INDICES=ri2)
>
> Anybody have ideas who knows about Histogram?

That is due to an array of all zeroes being passed to HISTOGRAM, with
MIN=1. It seems that *none* of your bins had a match in it, which is
very likely an indication that you need a larger search radius.

I've rolled the original algorithm into the function MATCH_2D,
available here:

http://turtle.as.arizona.edu/idl/match_2d.pro

This version deals gracefully with the case of no search points in any target point bin, and also returns -1 in all cases where no point was found within the search radius (vs. returning a close-but-not-necessarily-the-closest point).

In principle one could also write MATCH_ND, but the sensitivity to a poorly chosen search_radius would become quite extreme in dimensions higher than 2, as would the need to search 2^d adjacent bins.

JD
