Subject: Re: Filtering data in multidimensional arrays
Posted by davidf on Thu, 01 Jun 2000 07:00:00 GMT
View Forum Message <> Reply to Message

Simon de Vet (simon@mathstat.dal.ca) writes:

> David Fanning wrote:
>
>> Oh, the WHERE function applies equally well to multidimensional
>> arrays. And the beauty of it is, you don't have to understand
>> it. You just have to use it. :-)
>
> I'm clearly not capable of doing even that! :)
>
> I've followed the instructions on your site, but the results are not what I
> expected. Since the flights run from 3-19, I would expect that index to look
> like 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19. However, the actual index
> looks like this, but repeated many times, ending at #15. Weird...
>
> If I try to use my indices in the plotting, (ie: plot, pem1(flightindex,
> ptsindex, 4), pem1(flightindex,ptsindex,3)) IDL hangs. After about 5 minutes
> I cancel the process. Without culling negative values the process takes about
> 1.5s, at most.
>
> Any idea what could be going wrong? Am I misunderstanding the usage of these
> indices?

Humm, I'd have to see a little code to see what it is exactly you
are trying to do, but I have no time to look at it today. This is
a pretty small array, however. Why don't you just break the problem
up into (20?) pieces that you know how to deal with and do it in
a loop? It may cost you 0.000348302 seconds of processing time,
but it would be finished by the end of lunch today. :-)

Cheers,

David

--
David Fanning, Ph.D.
Fanning Software Consulting
Phone: 970-221-0438 E-Mail: davidf@dfanning.com
Coyote's Guide to IDL Programming: http://www.dfanning.com/
Toll-Free IDL Book Orders: 1-888-461-0155

Subject: Re: Filtering data in multidimensional arrays

Posted by Simon de Vet on Thu, 01 Jun 2000 07:00:00 GMT
View Forum Message <> Reply to Message

David Fanning wrote:

> Oh, the WHERE function applies equally well to multidimensional
> arrays. And the beauty of it is, you don't have to understand
> it. You just have to use it. :-)

I'm clearly not capable of doing even that! :)

I've followed the instructions on your site, but the results are not what I expected. Since the flights run from 3-19, I would expect that index to look like 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19. However, the actual index looks like this, but repeated many times, ending at #15. Weird...

If I try to use my indices in the plotting, (ie: plot, pem1(flightindex, ptsindex, 4), pem1(flightindex,ptsindex,3)) IDL hangs. After about 5 minutes I cancel the process. Without culling negative values the process takes about 1.5s, at most.

Any idea what could be going wrong? Am I misunderstanding the usage of these indices?

Simon

Subject: Re: Filtering data in multidimensional arrays
Posted by davidf on Thu, 01 Jun 2000 07:00:00 GMT
View Forum Message <> Reply to Message

Simon de Vet (simon@mathstat.dal.ca) writes:

> I have a 3 dimensional array, of size (20,40,13). This represents
> (Flight Number, Point Number, Data Type). The flights range from 3-19.
> Each flight has between 10 and 40 points. They all have the same number
> of data types. Using the replicate command, I have filled the empty
> array locations with -1111.
>
> Now, I want to plot this data, with all the negative values in the first
> two dimensions ('Bad Data') removed entirely. I have fooled around a bit
> with WHERE, but I can only understand it when it's applied to one
> dimensional arrays.
>
> How can I use it in my particular case? Is there another solution?

Oh, the WHERE function applies equally well to multidimensional

arrays. And the beauty of it is, you don't have to understand
it. You just have to use it. :-)

The secret is that you can use one-dimensional subscripts
to subset multi-dimensional arrays. If you really do want
to know how this works (and I wouldn't bother, probably)
here is an article on my web page that shows you how to
convert 1D subscripts of the type returned by WHERE to
multidimensional coordinates:

  http://www.dfanning.com/tips/where_to_2d.html

Cheers,

David

--
David Fanning, Ph.D.
Fanning Software Consulting
Phone: 970-221-0438 E-Mail: davidf@dfanning.com
Coyote's Guide to IDL Programming: http://www.dfanning.com/
Toll-Free IDL Book Orders: 1-888-461-0155

Subject: Re: Filtering data in multidimensional arrays
Posted by davidf on Fri, 02 Jun 2000 07:00:00 GMT
View Forum Message <> Reply to Message

Simon de Vet (simon@mathstat.dal.ca) writes:

> Perhaps I am misunderstanding the usage of the filter. Am I producing seperate
> filters for each dimension (hence the problem above), or am I producing a global
> filter for the entire array (which treats entries individually, and not genrically
> by dimension)?

You are producing a global filter for the entire array.
But I think Martin's code fragments point out how to
do this in a sensible way by handling only one dimension
at a time.

Cheers,

David

--
David Fanning, Ph.D.
Fanning Software Consulting
Phone: 970-221-0438 E-Mail: davidf@dfanning.com

Subject: Re: Filtering data in multidimensional arrays
Posted by Martin Schultz on Fri, 02 Jun 2000 07:00:00 GMT
View Forum Message <> Reply to Message

Hi Simon,

first of all: which data are you working on there: PEM-West or PEM-Tropics ?
Are you using merged data sets or are you working with the original files from the GTE website?

As to your problem: if it's only the plot you are concerned about, try using the MIN_VALID keyword and set it to a reasonably low number (usually 0. except for the longitude which may extend to -180 ;-). If you really need to get rid of that data (e.g. for computing a mean etc.), it depends on how you atore the data exactly. I usually have arrays with dimensions where observations loops from 0 to however many there were, and variables is the number of variables you read in. I also keep a string array containing the variable names. To distinguish between various flights, I have a variable named 'FLIGHT', and to have a continuous time vector, I add the julian day and the time (divided by 24.).
Then, say I want to filter a certain longitude range and eliminate all missing values, what I do is:

```
  ilon = where(StrUpCase(Names) eq 'LON')   ; if you go beyond your personal hack
                                  ; please add error checking !
  lon = reform( data )
  ok = Where(lon gt 150. OR lon le -100 )

  regional_data = data
  ; then I extract all variables I want to plot and eliminate missing values
  itime = where(StrUpCase(Names) eq 'CONTINUOUSTIME')
  io3   = where(StrUpCase(Names) eq 'OZONE')
  time = reform(regional_data)
  o3 = reform(regional_data)

  ok = where(o3 ge 0.)
  plot,time, o3
  print,'Mean, median of Ozone = ',mean(o3), median(o3)
```

Hope, this helps,
Martin

Simon de Vet wrote:
>
> David Fanning wrote:
>
>> Humm, I'd have to see a little code to see what it is exactly you
>> are trying to do, but I have no time to look at it today. This is
>> a pretty small array, however. Why don't you just break the problem
>> up into (20?) pieces that you know how to deal with and do it in
>> a loop? It may cost you 0.000348302 seconds of processing time,
>> but it would be finished by the end of lunch today. :-)
>
> I'm not even sure if I could do this, but I'll give it a try.. It's ackward since
> I'll have to eventually read in another 3 data sets with similar properties, and
> I'd much rather handle 4 big arrays than 80 little ones. It's my own brain I'm
> worried about, not the computer's!
>
> I think my problem may be that my array is not consistent. I can apply one filter
> to the flight# (3-19 only), but I cannot apply a single filter to the point
> number, since this varies from flight to flight. On flight 3, there may only be 10
> data points, but flight 4 may have 12, flight 4 21, etc...
>
> Perhaps I am misunderstanding the usage of the filter. Am I producing seperate
> filters for each dimension (hence the problem above), or am I producing a global
> filter for the entire array (which treats entries individually, and not genrically
> by dimension)?
>
> Simon

--
[[[[[[[[[[
[[Dr. Martin Schultz Max-Planck-Institut fuer Meteorologie [[
[[Bundesstr. 55, 20146 Hamburg [[
[[phone: +49 40 41173-308 [[
[[fax: +49 40 41173-298 [[
[[martin.schultz@dkrz.de [[
[[[[[[[[[[

Subject: Re: Filtering data in multidimensional arrays
Posted by Simon de Vet on Fri, 02 Jun 2000 07:00:00 GMT
View Forum Message <> Reply to Message

David Fanning wrote:

> Humm, I'd have to see a little code to see what it is exactly you
> are trying to do, but I have no time to look at it today. This is
> a pretty small array, however. Why don't you just break the problem
> up into (20?) pieces that you know how to deal with and do it in
> a loop? It may cost you 0.000348302 seconds of processing time,
> but it would be finished by the end of lunch today. :-)

I'm not even sure if I could do this, but I'll give it a try.. It's ackward since I'll have to eventually read in another 3 data sets with similar properties, and I'd much rather handle 4 big arrays than 80 little ones. It's my own brain I'm worried about, not the computer's!

I think my problem may be that my array is not consistent. I can apply one filter to the flight# (3-19 only), but I cannot apply a single filter to the point number, since this varies from flight to flight. On flight 3, there may only be 10 data points, but flight 4 may have 12, flight 4 21, etc...

Perhaps I am misunderstanding the usage of the filter. Am I producing seperate filters for each dimension (hence the problem above), or am I producing a global filter for the entire array (which treats entries individually, and not genrically by dimension)?

Simon
