Subject: Re: Subsetting an array with user defined values
Posted by Craig Markwardt on Mon, 19 Jul 2004 16:24:35 GMT
View Forum Message <> Reply to Message

dfinnegan@crrel.usace.army.mil (DCF) writes:

```
>  Hi all,
>
>  Not sure if this is a repetitive question or not.
>
>  I have a series of ascii XYX files that I am reading into IDL and
>  creating an array with (array[3,millions]). Importing is no problem as
>  well as working with the array. What I would like to do is subet the
>  array by a given min and max longitude and latitude so that I get back
>  only the values within that range, say a user defined box.
>
>  I am doing this with the following code (the data has been sorted and
>  nulls removed prior to this):
>
>  latmin=3763965.28
>  latmax=3763985.74
>  lonmin=534302.28
>  lonmax=534335.50
>
>  IF latmin GE min(data[0,*]) then begin
>        print, 'inside latmin'
>         vgood= where(data[0,*] GT latmin)
>         data =  temporary(data[*,vgood])
>  ENDIF
>
>  IF latmax LT max(data[0,*]) then begin
>        print, 'inside latmax'
>          vgood= where(data[0,*] LE latmax)
>          data = temporary(data[*,vgood])
>  ENDIF
>
>  IF lonmin GE min(data[1,*]) then begin
>          vgood= where(data[1,*] GE lonmin)
>          data =  temporary(data[*,vgood])
>  ENDIF
>
>  IF  lonmax LE max(data[1,*]) then begin
>          vgood= where(data[1,*] LE lonmax)
>          data = temporary(data[*,vgood])
>  ENDIF
>
>
>  Is there a more efficient way of extracting the data from these
```

> arrays? An example would be great!

Yes, probably. You can select the data with a single expression:

```
lat = data[0,*] & lon = data[1,*]
vgood = where(lat GE latmin AND lat LE latmax AND $
          lon GE lonmin AND lon LE lonmax)
```

You can't avoid computing selection expressions, but by computing them all at once, you will only extract your array once.

Also, if you are building up a selection expression, piece by piece, where you may not know how many expressions there are, then I like to accumulate a "mask" array, something like this:

```
mask = bytarr(n_elements(lon))
if lonmin NE -999 then mask = mask AND (lon GE lonmin)
if lonmax NE -999 then mask = mask AND (lon LE lonmax)
if latmin NE -999 then mask = mask AND (lat GE latmin)
if latmax NE -999 then mask = mask AND (lat LE latmax)

vgood = where(mask)
```

This allows you to set the min/max values to -999 to disable them if desired, but still has the efficiency of extracting the array only once.

Of course, you also need to check the return value(s) of WHERE to be sure that you haven't selected a zero-length array.

Good luck,
Craig

--
-- ----------
Craig B. Markwardt, Ph.D. EMAIL: craigmnet@REMOVEcow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response
-- ----------