Subject: Re: fastest way to find the first non-zero value in an array
Posted by Clemens on Tue, 08 Apr 2008 08:42:50 GMT
View Forum Message <> Reply to Message

it depends on the size of array. Small array is a simple loop.
Huge one can be :

index_arr = where(arr NE 0)
print, index_arr[0]

on very huge arrays you can devide arr into big peaces.

Clemens

<smasson@locean-ipsl.upmc.fr> schrieb im Newsbeitrag
news:ff3bf651-8707-47e8-8cff-5e694dd5622f@m71g2000hse.google groups.com...
> Hi,
>
> I want to find the first non-zero value of an array. Is there a faster
> way to do this than with the "where" command: (where(array ne 0))[0]
> "Where" will look for all non-zero values and I only need the first
> one. It would be great if I could stop "where" in its search process
> as soon as it found one element...
>
> sebastien

Subject: Re: fastest way to find the first non-zero value in an array
Posted by Spon on Tue, 08 Apr 2008 13:03:51 GMT
View Forum Message <> Reply to Message

On Apr 8, 9:04 am, smas...@locean-ipsl.upmc.fr wrote:
> Hi,
>
> I want to find the first non-zero value of an array. Is there a faster
> way to do this than with the "where" command: (where(array ne 0))[0]
> "Where" will look for all non-zero values and I only need the first
> one. It would be great if I could stop "where" in its search process
> as soon as it found one element...
>
> sebastien

Hi Sebastien,

I can think of two separate ways of going about it:

Firstly, for certain arrays it may be perfectly sensible to use the
'inefficient', FORTRAN 101 route:

```
i = 0
WHILE Array[i] EQ 0 DO i++
RETURN, i
```

Secondly, I've attempted to do it using HISTOGRAM below.

```
 ; Is the first element non-zero?
 ; If it is, we can save ourselves a lot of
 ; hassle...
IF Array[0] NE 0 THEN RETURN, 0L

 ; We now know element 0 contains data = 0.
 ; Generate a histogram of the array such
 ; that this element will always be put in the
 ; first bin:
H = HISTOGRAM(CEIL(ABS(Array)), REVERSE_INDICES = RI)

 ; Are there non-zero elements in the array?
IF N_ELEMENTS(H) EQ 1 THEN MESSAGE, $
  'Array contains only zeroes!'

 ; Array[0] is always going to be in the first bin.
 ; Get the contents of that bin:
BinContents = RI[RI[0]:RI[1]-1]

 ; How many drops in that bin?
NBC = N_ELEMENTS(BinContents)

 ; Where do the drop indices stop increasing linearly?
 ; That's where the first non-zero element must be.
Index = WHERE((LINDGEN(NBC) - BinContents) NE 0, Count)

 ; If all the zeroes in the array come before the first
 ; non-zero value, then we won't get any indices returned,
 ; but the next index will be the next element after the
 ; end of our BinContents vector. This must be non-zero.
IF Count EQ 0 THEN RETURN, NBC

 ; Otherwise, use the WHERE results to return
 ; the index of first non-zero element of the array.
RETURN, Index[0]
```

As far as I can tell, both methods work; and which is going to be faster (between these two and just using WHERE) is going to depend on your array.

Let us know how you get on,

Regards,

Chris

Subject: Re: fastest way to find the first non-zero value in an array
Posted by Vince Hradil on Tue, 08 Apr 2008 13:15:22 GMT
View Forum Message <> Reply to Message

On Apr 8, 3:04 am, smas...@locean-ipsl.upmc.fr wrote:
> Hi,
>
> I want to find the first non-zero value of an array. Is there a faster
> way to do this than with the "where" command: (where(array ne 0))[0]
> "Where" will look for all non-zero values and I only need the first
> one. It would be great if I could stop "where" in its search process
> as soon as it found one element...
>
> sebastien

Wow - I'd be interested in knowing how slow 'where' is. Are we
talking the difference between 0.01 seconds and 0.05 seconds? Or even
the difference between 1 and 5 seconds? Time is money, but at what
point does our 'need for speed' end?

Subject: Re: fastest way to find the first non-zero value in an array
Posted by David Fanning on Tue, 08 Apr 2008 13:34:38 GMT
View Forum Message <> Reply to Message

Vince Hradil writes:

> Wow - I'd be interested in knowing how slow 'where' is. Are we
> talking the difference between 0.01 seconds and 0.05 seconds? Or even
> the difference between 1 and 5 seconds? Time is money, but at what
> point does our 'need for speed' end?

I've gotten to the point where anything that takes less time
than it takes to go get a cup of coffee is fast enough. I
used to think fast, elegant programs were required. But
when you are writing one-offs day after day, why bother?
With Starbucks just across the street, I can afford to be
a little loose with a FOR loop.

Cheers,

David

P.S. That said, I just spent the entire weekend re-working
a program I inherited from someone else. It is generally a
good idea to write a program in such a way that someone else
can get it to work in less time than it takes to write the
darn thing from scratch. :-)

--
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: fastest way to find the first non-zero value in an array
Posted by Spon on Tue, 08 Apr 2008 13:56:33 GMT
View Forum Message <> Reply to Message

On Apr 8, 2:15 pm, Vince Hradil <hrad...@yahoo.com> wrote:
> On Apr 8, 3:04 am, smas...@locean-ipsl.upmc.fr wrote:
>
>> Hi,
>
>> I want to find the first non-zero value of an array. Is there a faster
>> way to do this than with the "where" command: (where(array ne 0))[0]
>> "Where" will look for all non-zero values and I only need the first
>> one. It would be great if I could stop "where" in its search process
>> as soon as it found one element...
>
>> sebastien
>
> Wow - I'd be interested in knowing how slow 'where' is. Are we
> talking the difference between 0.01 seconds and 0.05 seconds? Or even
> the difference between 1 and 5 seconds? Time is money, but at what
> point does our 'need for speed' end?

You make a good point, Vince.

If anyone can find an array that IDL actually searches faster using my
Histogram-and-small-Where setup than by using "Where" on its own, let
me know! :-)

Chris

Subject: Re: fastest way to find the first non-zero value in an array
Posted by Vince Hradil on Tue, 08 Apr 2008 14:02:38 GMT

On Apr 8, 8:34 am, David Fanning <n...@dfanning.com> wrote:
> Vince Hradil writes:
>> Wow - I'd be interested in knowing how slow 'where' is. Are we
>> talking the difference between 0.01 seconds and 0.05 seconds? Or even
>> the difference between 1 and 5 seconds? Time is money, but at what
>> point does our 'need for speed' end?
>
> I've gotten to the point where anything that takes less time
> than it takes to go get a cup of coffee is fast enough. I
> used to think fast, elegant programs were required. But
> when you are writing one-offs day after day, why bother?
> With Starbucks just across the street, I can afford to be
> a little loose with a FOR loop.
>
> Cheers,
>
> David
>
> P.S. That said, I just spent the entire weekend re-working
> a program I inherited from someone else. It is generally a
> good idea to write a program in such a way that someone else
> can get it to work in less time than it takes to write the
> darn thing from scratch. :-)
>
> --
> 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.")

Exactly - the google-mentality is making everyone think that 10 seconds is too long to wait for anything. But this makes sense to me: it take about 2-4 hours for the scientist to acquire the data for a certain experiment, it takes me about the same time to create the one-off and run the analysis for said experiment. Sure I could spend about 8 hours to tweak the analysis to make it run in 5 minutes instead of half-an-hour, but why bother. Unless, of course we need that extra half-hour? Then I end up writing some obfuscated code (http://en.wikipedia.org/wiki/Obfuscated_code) that uses histograms, and the next developer that comes along just says, 'unh?' and re-writes the whole thing.

Subject: Re: fastest way to find the first non-zero value in an array
Posted by karo03de on Wed, 09 Apr 2008 20:02:32 GMT

On 8 Apr., 10:42, "Clemens" <yy...@hotmail.com> wrote:
> it depends on the size of array. Small array is a simple loop.
> Huge one can be :
>
> index_arr = where(arr NE 0)
> print, index_arr[0]
>
> on very huge arrays you can devide arr into big peaces.
>
> Clemens
>
> <smas...@locean-ipsl.upmc.fr> schrieb im
Newsbeitragnews:ff3bf651-8707-47e8-8cff-5e694dd5622f@m71g200 0hse.googlegroups.com...
>
>> Hi,
>
>> I want to find the first non-zero value of an array. Is there a faster
>> way to do this than with the "where" command: (where(array ne 0))[0]
>> "Where" will look for all non-zero values and I only need the first
>> one. It would be great if I could stop "where" in its search process
>> as soon as it found one element...
>
>> sebastien

What about the system function ARRAY_EQUAL ?
Karsten
