
Subject: Definition of Median was(Re: Finding the index of the median)

Posted by [Thomas A. McGlynn](#) on Wed, 30 Oct 1996 08:00:00 GMT

[View Forum Message](#) <> [Reply to Message](#)

In looking at the get the index of the median value, I noted that the behavior of the IDL median filter is not what I would have expected.

E.g.,

```
print, median([1,2,3,10])
```

prints out 3. This is independent of the order of elements in the array. Is there an accepted definition of what the median value is in this case. For example, I might think 2.5 is a more appropriate choice (but one which would have made the previous discussion incorrect).

Perhaps more important, if the selection of 3 rather than 2 as the median arises because $3 > 2$, then does the IDL median function systematically 'overestimate' the median. More precisely given two values that are reasonable choices for the median does IDL always choose the later?

I don't like IDL's choice here. In some of my work I've used a n-dimensional median definition as the point (or points) at which the unit to all the array elements sum to zero. This gives a nice generalization of the concept of median to higher dimensions but it would require a value (any value) m where $2 < m < 3$ in the current instance.

Tom McGlynn
tam@silks.gsfc.nasa.gov

Subject: Re: Definition of Median was(Re: Finding the index of the median)

Posted by [kgb](#) on Tue, 05 Nov 1996 08:00:00 GMT

[View Forum Message](#) <> [Reply to Message](#)

In article <31OCT199600081707@sorbet.gsfc.nasa.gov>, landsman@sorbet.gsfc.nasa.gov (Wayne Landsman (301)-286-3625) writes:

> In article <E0402F.Ao8@midway.uchicago.edu>, meron@cars3.uchicago.edu writes...

>> In article <327771BA.2781@silks.gsfc.nasa.gov>, "Thomas A. McGlynn" >

> Now the median is a nonlinear filter so "flux conservation" isn't necessarily

> an important consideration. In any case, for smoothing it always makes

> sense to use an odd value of the Width parameter. But I had to write a

> program to median combine a stack of N images (combining CCD flatfield images

> while removing cosmic rays), and it took me a long time to figure out why

> my flatfield normalization was screwed up whenever N was even.
> The normalization problem was fixed when I stopped using the MEDIAN() function
> but instead SORT()'ed the values, and then averaged the two values on either
> side of the median dividing line. (In case anyone is interested, my median
> stacking program is available at
>
> <http://idlastro.gsfc.nasa.gov/ftp/pro/image/medarr.pro>)

Be careful with this too. If you have say 4 images and a pixel where there is a cosmic ray in two of them the median ends up being nowhere near any of the data values. If you use the median as a first pass estimate before doing cr rejection you can end up rejecting ALL the values!

I once had this problem with HST data...

Karl

--

kgb@aaoepp.aao.gov.au [Anglo-Australian Observatory]

----> pubs: <http://www.ast.cam.ac.uk/~kgb/papers.html>

----> pgperl: <http://www.ast.cam.ac.uk/~kgb/pgperl.html>
