Subject: Re: Mode function for floating point arrays
Posted by Rob Klooster on Mon, 22 Jul 2013 11:24:49 GMT
View Forum Message <> Reply to Message

Op zaterdag 20 juli 2013 01:50:08 UTC+2 schreef Matthew Argall het volgende:
> It seems like the "goodness" of this lies in how well the UNIQ function can determine if two numbers are truly unique. Then, after that, how well Value_Locate can match unique values to their duplicates. Is that right?

Exactly, UNIQ() is used for comparing floats to see if they are equal or not. You could change some lines in that function from:
indices = where(q ne shift(q,-1), count)
to:
indices = where(abs(q - shift(q,-1)) gt eps, count)
for a fixed value of eps. Be careful with this kind of comparisons, as the value to take for eps is not well defined. Take a look at this article which explains all the pitfalls when comparing floating point numbers:

http://www.cygnus-software.com/papers/comparingfloats/compar ingfloats.htm

Value_locate() will work whatever the input is, since it does not look at exact matches. It will just find the interval to which a specific number belongs. You just need to make sure that the UNIQ() function outputs the lowest number of a particular bin.

>> Note that two floats are only assumed equal when they have the exact same binary value.
>
>
>
> I think there is more information in this sentence than I can grasp at the moment... Is there any reason to suspect that the precision of the result is less than the precision of the numeric type of the input array?

Again, have a look at the article. It will make things a bit clearer.

Rob.