
Subject: Re: remove duplicates WITHOUT sorting
Posted by [JD Smith](#) on Mon, 12 Feb 2007 21:15:41 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Mon, 12 Feb 2007 09:04:45 -0800, rpertaub@gmail.com wrote:

> Hello All,
> I have an array of 2xn which is a list of coordinates (x,y positions)
> and I need to remove any duplicates in this array. Since they are x,y
> coordinates I cannot really sort the array, hence cannot use the Uniq
> function. Is there another way of doing this?

It's easiest to recast as 1D. HIST_2D (or HIST_ND) does this for you, but
it's easy to do yourself:

$\text{index} = x + (\max(x) + 1) * y$

and then using UNIQ on this list of indices should give you the row
positions of unique coordinates. HISTOGRAM can work as well (either
with HIST_ND, or by first constructing this index vector above), and
it will be faster, but, as usual, will consume lots of memory (and
potentially be very slow) if your coordinates are sparsely sprinkled over
a large range of values (your current example is somewhat sparse, but not
horrible). UNIQ, with its SORT based implementation, doesn't suffer from
that issue.

JD
