Subject: Re: Check for duplicate locations
Posted by Craig Markwardt on Wed, 24 Feb 2016 14:14:41 GMT
View Forum Message <> Reply to Message

On Thursday, February 18, 2016 at 11:48:38 AM UTC-5, Med Bennett wrote:
> I have X,Y,Z data for several thousand points that I need to check for duplicate locations. I cannot have duplicated locations in the sample data, as it breaks the kriging algorithm I am using. I've always used a brute force method of computing a distance function between each point and all subsequent points, and flagging any points for which the distance is zero, or some small threshold. This method is very slow for larger numbers of points, however. Does anyone have a method for doing this more efficiently? I've found simple methods for one-dimensional data, but not for points in 3-space.
>
> Thanks!

There's no simple answer to this. Usually you need to do some kind of filtering to make it a faster process. For example, if the spread in "Z" values in your 3-space is the greatest, sort by Z first, then you can limit the range over which you perform the 3D distance computation. But there are lots of potential gotchas when you do this! Another solution is to get a faster CPU!

Craig